# INSTRUCTIONAL COACHING EVALUATION USING A REALIST APPROACH: A MIXED METHODS STUDY

**By**
**Claudia Berlage**
**Lic. Oec. Publ., University of Zurich, Switzerland, 1994**

**A dissertation submitted in partial fulfillment of the requirements**
**For the degree of Doctor of Education**
**in the Department of Educational Leadership**
**in the College of Education**
**at the University of Bridgeport,**
**Bridgeport, CT**
**Fall Term**
**2015**

University of Bridgeport
Committee Approval of a Dissertation

Dissertation Title:
Instructional Coaching Program Evaluation Using a Realist Approach: A Mixed Methods Study

Submitted by: Claudia Berlage

I have read this dissertation and have found it to be of satisfactory quality for a doctoral degree.

11/23/15
11/23/2015

Thomas W. Christ, Ph.D.
Chair Dissertation Committee

I have read this dissertation and have found it to be of satisfactory quality for a doctoral degree.

11/23/2015
11/23/2015

Gail Prelli, Ed.D.
Member, Dissertation Committee

I have read this dissertation and have found it to be of satisfactory quality for a doctoral degree.

11/23/2015
11/23/2015

James Ritchie, Ph.D.
Member, Dissertation Committee

I have read this dissertation and have found it to be of satisfactory quality for a doctoral degree.

11/23/15
11/23/2015

Thomas W. Christ. Ph.D.
Director Education Leadership Doctoral Program

I have read this dissertation and have found it to be of satisfactory quality for a doctoral degree.

11/23/2015
11/23/2015

Allen Cook, Ph.D.
Dean of the School of Education and
Member, Dissertation Committee

# Abstract

The necessity to evaluate professional development has become more pressing for districts in a time of increased accountability. A number of empirical studies had shown that it was difficult to measure the effects of professional development, especially its effect on transferring practice and student achievement.

The purpose of this study was to evaluate instructional coaching professional development. A realist evaluation approach was used to develop a realist evaluation based framework and to pilot the instruments to address the need as determined by the literature to effectively evaluate instructional coaching.

This study was an evaluation and as such a mixed method study, conducted at one comprehensive high school with instructional coaches, Algebra teachers and students as participants. The researcher a teacher pre-and post survey at the beginning and at the end of the school year, a student survey at the end of the school year as well as a coach survey with the teachers at the end of the year,. During the school year teachers and coaches wrote learning log entries on a quarterly basis and at the end of the year the researcher conducted semi structured interviews with all the participating teachers. All data was able to address five levels of professional development evaluation: teachers' reactions, teachers' learning, organizational support, teachers' use of new knowledge and skills, and students' learning. Teachers perceived that the instructional coaching increased their knowledge about and their use of instructional practices such as the implementation of group work and the provision of more student centered activities to engage students and promote student to student discourse. Teachers perceived their students to demonstrate increased collaboration in groups, more perseverance, and more independence. The

knowledge about the asking of higher- level questions that require students to explain their thinking as well as the provision of multiple representations to support visualization of skills and concepts, showed a significant difference between the average pre and post test scores. The instructional coaching professional development also lead to an increased teacher use of consistently asking higher level questions that require students to explain their thinking as indicated by a significant difference between the average pre and post test scores. The purpose of a realist evaluation was to discover what it was about a program that worked for whom and in what circumstances. The results of the study indicated that the context in which the professional development took place and the perceptions of the teachers about the coaches made a difference in how beneficial the teachers perceived the professional development to be. The realist approach to evaluate instructional coaching professional development is an appropriate design to consider due to the contextual nature of coaching. It provided a process approach to causality with the intent to measure what mechanisms lead to what outcomes and why, while constructing the measures according to theory based assumptions about the causal path of professional development.

## Dedication

**For my family**

**Karsten, Kristina, Tessa, Antonia, and Grace**

**and my parents**

**Horst and Maria Dittrich**

*Now unto him that is able to do exceeding abundantly above all that we can ask or think. Ephesians 3.20*

# Table of Contents

# List of Tables

# List of Figures

## Chapter 1. INTRODUCTION

School districts' interest in evaluation of professional development had grown as a result of the increased pressure of accountability (Guskey, 2000). Major criticisms of evaluations of professional development were that they were not evaluations but mere documentations, that they were too shallow, and too brief (Guskey, 2000). Quality in accountability and assessment was missing from many professional development efforts (Reeves, 2010).

After interviewing more than a 150 teachers across the United States, Knight (2007) concluded: "Teachers do not resist change so much as they resist poorly designed change initiatives" (p.3). Traditional professional development appears to not be effective (Elmore, 2004; Joyce & Showers, 2002; Reeves 2010). Bush's study (Knight, 2007) showed a maximum of 10% implementation rate of traditional professional development. Knight noted from experience that more than 90% of teachers embraced and implemented programs that improved students' experiences in the classroom, when they received an appropriate amount of support. His research had grown into the promotion of an approach referred to as 'instructional coaching'. According to Fullan (2007), student learning depended on continual teacher learning, teacher collaboration, and the role of teacher's working context to facilitate teacher learning, all of which instructional coaching provides.

Gulamhussein (2013) noted that the Common Core State Standards focused on critical thinking and therefore professional development needed to emphasize practices that would turn students into critical thinkers and problem solvers. The most significant cost for districts to provide effective professional development that supported teachers in implementing Common Core State Standards would be purchasing time for teachers to

spend with coaches (Gulamhussein, 2013).

The school district targeted for this study started to gradually implement Mathematics and English Language Arts Common Core Curriculum since the 2010-2011 school year as the State of Connecticut adopted those standards. As part of a three-year implementation plan, teachers had been trained and had attended Common Core conferences hosted through the General Electric Foundation. Professional development focused on introducing the shifts in English Language Arts, and Mathematics, conveying to the participants the coherence of the standards through the grades and clarifying rigor and focus. The district had chosen a classroom embedded instructional coaching model to support their teachers to implement Common Core Standards for Mathematical Practice and to provide professional development in instructional practices that supported the implementation.

It was essential to evaluate the coaching program and to ensure that the evaluation was rigorous and ongoing using multiple measures (Killion, Harrison, Bryan, & Clifton, 2012). It was difficult to isolate the effects of instructional coaching (APQC Education Advanced Working Group, (2011)). Therefore a realist approach (Maxwell, 2012; Pawson & Tilley, 1997) was used to develop a framework that effectively evaluated instructional coaching professional development.

The framework was based on Guskey's (2000) model for evaluation of professional development and a logic model based on Desimone's (2009) core action theory. The process included the development and validation of instruments that were constructed aligning instructional practices that manifested implementation of the Common Core State Standards for Mathematical Practice in mathematics classrooms

with Guskey's (2000) levels of evaluation. The realism perspective was used to explain the mechanisms and measure their effects on outcome patterns.

**Problem Statement**

The district selected for this study became part of the Alliance District Program in Connecticut in 2013 that sought to close the achievement gap by turning around persistently low performing schools. The Alliance District Improvement Plan (ADIP 2013-2014) focused on the School Improvement Plans (SIP's) to promote the implementation of Common Core State Standards by aligning curriculum, changing pedagogy, and reducing the achievement gap.

The Alliance District Improvement Plan listed professional development activities as a priority. The purpose of these activities was to change pedagogy to facilitate implementation of the Common Core State Standards as a means to increase student achievement.

The district chose a classroom embedded instructional coaching model. Joyce and Showers' (2002) seminal study showed the effect of coaching as part of the professional development design on transfer of instructional practices (an increase from 5% to 95%). The literature indicated that for professional development to be successful, it must be ongoing, job embedded, interactive, and integrated (Desimone, 2009; Fogarty & Pete, 2007; Joyce & Showers, 2002; Tallerico, 2005).

The process in the district, where the study was conducted, involved an instructional coach and a teacher on special assignment (TOSA) visiting Algebra classrooms up to 11 times for each teacher throughout the school year and providing

constructive oral feedback and suggestions of what to work on for next time. The teachers practiced implementing suggestions in between visits (Classroom Embedded Professional development Protocol, Stamford Public Schools, 2014a).

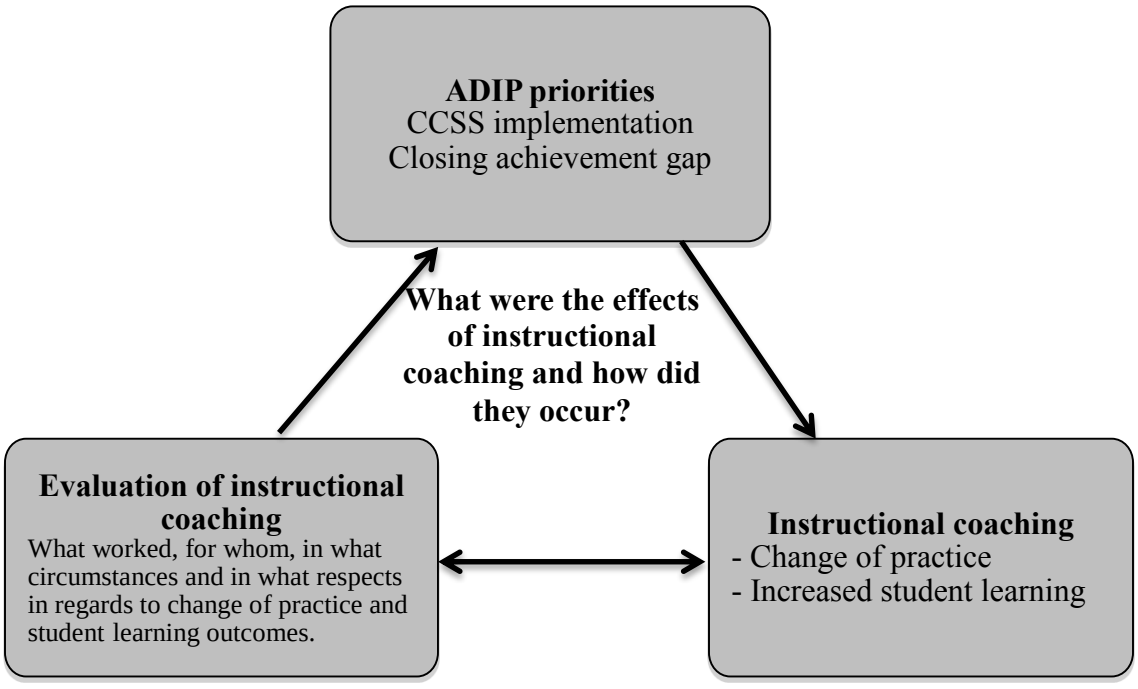
For the district to know if instructional coaching had reached its goals, it was necessary to evaluate the professional development. Research had shown that it was difficult to measure the effects of professional development, especially its effect on transferring practice and student achievement (Desimone, 2009; Garet et al., 2010; Guskey, 2000; Joyce & Showers, 2002; Reeves, 2010). The final report of Best Practices In Evaluating the Impact of Instructional Coaching on Teaching and Learning (APQC, 2011) concluded that it was challenging to isolate the results of coaching but that all districts reported significant improvement, linked to the instructional coaching program, and the use of multiple instruments to measure the impact of the intervention. For example, Killion, Harrison, Bryan, and Clifton (2012) recommended that central office administrators plan and implement a coaching program evaluation and to create evaluation instruments that are aligned with the program's goals.

The realist approach provided a unique methodology for evaluation whose distinctive viewpoint was not to measure the effect of instructional coaching with a 'regularity conception of causality' (Maxwell, 2004), but applying a 'process approach to causation' to determine 'what worked for whom in what circumstances and in what respects' (Maxwell, 2012; Pawson & Tilley, 2004). This study provided guidance to district leadership, building administration as well as the teachers, with an evaluation for their professional development to determine if their practice had changed, and why and how it had changed. It also measured to what degree instructional strategies that

implemented the Standards of Mathematical Practice were used, and if and how learning outcomes had been affected to inform the further refinement of the professional development offerings in the district.

Figure 1 conceptualized the relationships between the overarching goal-setting in the Alliance District Improvement Plan, the design of the professional development, and the contribution of the evaluation of the instructional coaching. The Alliance District Improvement Plan goals drove the design of the professional development and its focus. The instructional coaching professional development determined the type of evaluation that was necessary to examine the effects of the intervention towards the overarching goals: the change in teaching practice and in student learning outcomes. The outcome of the evaluation informed the future design of the professional development based on what worked for whom in what circumstances and in what respect.

Figure 1. Problem Statement Model



This study developed and piloted multiple instruments that were evaluated using a realist approach to describe the mechanisms, context, and outcome patterns of the instructional coaching. The seven instruments used in this study were: (1) Instructional coach survey; (2) Semi-structured interview questions; (3) Pre and post survey for teachers; (4) Teacher learning log rubric; (5) Classroom visitation protocol; (6) Mathematics classroom observation guide; and (7) Student survey instrument. The consistent alignment of all instruments to the Mathematics classroom observation guide (6) that was developed by General Electric allowed a focused evaluation of the professional development and its outcome patterns regarding the implementation of Common Core State Standards for Mathematical Practice. Furthermore all instruments were designed to address several of Guskey's (2000) levels to ensure multiple measures for each level. This comprehensive framework to evaluate instructional coaching should be replicable for other sites, different content, and in other districts.

**Purpose of the Study**

The purpose of this study was to use a realist approach to develop a theory based evaluation framework and to create and pilot the instruments to evaluate the instructional coaching for mathematics. The instruments were developed in such a way that intervention measures were aligned to theory based potential mechanisms in the context of the instructional coaching professional development to explain and measure the effect on desired outcome patterns (teacher learning, teachers' use of new knowledge and skills, and student learning outcomes). The evaluation framework for instructional coaching

models could be utilized for other contents and sites. The study was intended for the district leadership and administration as well as the participating teachers to gain information regarding the effectiveness of the professional development and to inform further adjustment and refinement.

Ultimately the study was adding to the body of knowledge to evaluate instructional coaching professional development through the development and piloting of a realist evaluation framework. The framework was replicable due to its research based concise format of the instruments and the alignment with Standards for Mathematical Practice based instructional strategies for mathematics classrooms.

**Role of Theory**

According to Neuman (2006) theory framed the way we look at or think about a topic. Theory provided concepts, basic assumptions, guided us to the important questions, and suggested ways to make sense of data. Without theory the researcher was prone to collect useless data, think vaguely and undermine the research effort resulting in a lack of focus.

On a macro level, the role of theory within this study was defined by the worldview and research paradigm of the researcher. The underlying worldview of this study was that of a realist perspective. This assumed a testable reality (theory was confirmed) with contextual explanation (theory was constructed) (Maxwell & Mittapalli, 2010).

Organizational theory and human performance technology (HPT) built the overarching conceptual frameworks for the study. The viewpoint of ‘realist evaluation’

guided the research design as a whole and the development of the instruments (Pawson & Tilley, 1997, 2004). Program evaluation literature provided the theoretical concepts and tools to frame the evaluation instruments within a logic model (Hatch & Cunliffe, 2013; Pershing, 2006; Wholey, Hatry & Newcomer, 2010).

Guskey's (2000) framework for professional development evaluation was specifically setting the evaluation within the context of professional development. His framework was used to develop the instruments, by which the instructional coaching was evaluated. His evaluation framework also defined the logic of outcome patterns the instruments were trying to explain and measure.

Joyce's and Showers' (2002) research regarding the design of professional development and how to promote knowledge transfer into used practice as well as student achievement was the theoretical foundation to define the criteria and conditions for an effective professional development. It was primarily used to justify the chosen professional development design**.** The program evaluation questions were framed by these studies.

Maxwell (2013) defined theory as a set of concepts and ideas and the proposed relationships among these, as "a structure that was intended to capture or model something about the world" (p.48). Maxwell charged the researcher to consider pre-existing theories and to understand the limitation of their use. He cautioned that uncritical use of theory threatened not only the credibility of the findings but also the ability of the research to contribute to further understanding. Maxwell (2013) pointed out that qualitative researchers fail to use existing theory mainly in two ways: they either do not use it enough, or they used it too uncritically.

In exploratory research the role of theory was to provide a framework and insights in the studied phenomenon without imposing theory on the study design but guiding further decisions. Theory was continually constructed through constant interaction with the data (Maxwell, 2013).

In explanatory research, the role of theory was to provide a framework that identified assumptions and concepts that could then be related through the study and provide causal or structural explanations for the phenomena. In explanatory research it depended on the phenomenon studied. If there was a lot of prior literature and research, existing theory could have a greater influence and the existing concepts could be the guide to develop research questions and even a research design. However, the limitations and underlying assumptions needed to be identified so the researcher stayed open to other and new explanations (Maxwell, 2013). In confirmatory research, theory provided the framework to develop a hypothesis that could be tested to further contribute to the theory or reject it (Maxwell, 2013).

This study contained confirmatory research in piloting instruments based on existing theory (Desimone, 2009; Guskey, 2000; Maxwell, 2012; Pawson, 2006, 2013) to evaluate a particular professional development design in a systematic, replicable way. Joyce and Shower's (2002) work as well as other literature regarding professional development design, were the framework that the evaluation questions were based on to determine the effects of the intervention.

This study contained explanatory research through qualitative methods that sought to provide additional contextual explanations for the findings to explain what

worked for whom in what circumstances and in what respect (Maxwell, 2012; Pawson, 2013; Pawson & Tilley, 1997).

The study also contained exploratory research through the development of a theory based evaluation framework (Desimone, 2009; Guskey, 2000; Pawson & Tilley, 2006; Joyce & Showers, 2002) that was piloted and so contributed to the body of knowledge regarding 'realist evaluation' to provide causal explanations in contextual settings (Maxwell, 2004, 2012; Maxwell & Mittapalli, 2010; Pawson, 2006).

## Research Questions

The evaluation questions were aligned to the problem statement, that the district, administration, and teachers needed to determine the effects of the instructional coaching professional development on teacher practice and student learning outcomes and to gain information how to refine the professional development for the following school year. The overarching research questions addressed the purpose of the study, to utilize a 'realist evaluation' approach to evaluate professional development and to determine if it has potential as a replicable framework.

*Overarching research questions:*

1. Was it possible to create a 'realist evaluation' based framework to evaluate an instructional coaching model of professional development?
2. Realist Evaluation: What context-mechanism-outcome pattern configurations did the instructional coaching confirm to inform the further refinement of the professional development?

3. Exploratory Factor Analysis: Did the student survey measure teacher practice and student learning outcomes regarding the implementation of the Common Core State Standards for Mathematical Practice?

The professional development evaluation questions were based on the four conditions Joyce and Showers (2002) listed as essential for staff development to significantly affect student learning. This was linked to Guskey's (2000) five levels framework for professional development evaluation. First of all it required a community of professionals that came together, who studied together, put into practice what they were learning, and shared the results. Next, the content of staff development was formed around curricular and instructional strategies selected because they had a high probability of affecting student learning and the students' ability to learn. In addition the magnitude of change generated was sufficient that changes in the student's knowledge and skills were palpable. What was taught, how it was taught, and the social climate of the school had to change to the degree that the increase in student ability to learn was manifested. Lastly, the processes had to enable educators to develop the skills to implement what they were learning (Joyce & Showers, 2002).

Guskey's (2000) five levels framework for professional development evaluation was aligned with the evaluation questions in this study as follows:

1. How did the teachers describe their experience with the instructional coaching professional development (Guskey's levels 1 and 3)?
2. In what way did the instructional coaching professional development have an effect upon teachers' perceptions of their knowledge about teaching practices that

implement the Common Core State Standards for Mathematical Practice (Guskey's level 2)

3. In what way did the classroom instructional coaching professional development have an effect upon teachers' perceptions of their use of teaching practices that implement Common Core State Standards for Mathematical Practice (Guskey's level 4)?
4. In what way did the instructional coaching professional development have an effect upon teachers' perceptions of their students' learning defined as the demonstration of skills that evidence the implementation of Common Core State Standards for Mathematical Practice in a classroom (Guskey's level 5)?
5. Teachers and students reported that instructional coaching professional development affected their use of teaching practices that implement the Common Core State Standards for Mathematical Practice (Guskey's level 4).
6. Teachers and students reported that instructional coaching professional development affected students' learning outcomes as measured by the demonstration of skills that evidence the implementation of Common Core State Standards for Mathematical Practice (Guskey's level 5)

**Significance of the Study**

This study was significant for three main reasons: It evaluated the instructional coaching professional development which according to Killion et al. (2012) provided necessary information to the district leadership to refine the professional development. The evaluation of the instructional coaching professional development served as the

vehicle to develop and pilot a 'realist evaluation' based framework to evaluate instructional coaching, identifying and explaining different effects of the professional development organized by Guskey's (2000) levels of evaluation. Teacher knowledge, teacher transfer (change in practice), and student learning were used to organize outcome patterns to address the challenges of evaluating instructional coaching as stated by the final report of Best Practices In Evaluating the Impact of Instructional Coaching on Teaching and Learning (APQC, 2011). ). Finally, this study contributed to the field of Mixed Methods Research and supported Maxwell's (2004) argument that qualitative research provides causal explanations in contextual settings as well as the confirmation and application of a 'realist evaluation' approach (Pawson, 2006, 2013) to discover what it is about a program that worked, for whom, in what circumstances and why.

**Limitations of the Study**

This study was a case study and hence bound by a specific context and case (Yin, 2009). This study was limited to one high school in one urban/suburban district and to one department (Mathematics) that received instructional coaching professional development with two coaches, one for Algebra 1, and one for Algebra 2, and the twelve teachers participating. The evaluation framework used in this study is likely to be applicable to other sites that intend to implement Common Core State Standards utilizing instructional coaches. However the results obtained in this study were limited to the perspectives of the selected group of mathematics teachers participating in this study, and was not representative of other teachers in the school or in the district. The knowledge gained from this study supports the use of a realist perspective applied to professional

development evaluation theory and is expected to be applicable to other subject areas and settings and applied to settings outside of this case.

**Definition of Terms**

**Instructional Coaches:** Instructional coaches collaborate with teachers to support them to incorporate research-based instructional practices (Knight, 2007).

**Professional development:** According to Showers, Joyce, and Bennett (1987) the purpose of professional development design is to create the conditions, which allow sufficient levels of knowledge and skills to be developed to sustain practice and to provide the condition to support practice until executive control has been achieved and transfer has occurred.

**Professional Development Evaluation:** Evaluation is a systemic investigation to measure merit or worth (Guskey, 2000). It should therefore be results- or goal driven.

**Program:** "A program is a set of resources and activities directed toward one or more common goals, typically under the direction of a single manager or management team" (Newcomer, Hatry, & Wholey, 2010, p.49).

**Program Evaluation:** Program evaluation according to Newcomer et al. (2010) is the application of systematic methods to determine program operation outcomes.

## Chapter 2: REVIEW OF THE LITERATURE

This chapter reviewed the literature to date pertaining to professional development evaluation. Special attention was given to the instructional coaching model of professional development to create a framework to effectively evaluate this model. Organizational theory served as the overarching lens through which the view of the organization as a system was established (Hatch & Cunliffe, 2013; Senge, 2006).

Human Performance Technology (HPT) literature and Program Evaluation viewed the programs within the organization systematically and results oriented with the purpose of measuring and improving performance (Pershing 2006). The development and significance of instructional design and development models such as 'ADDIE' (Analyze, Design, Development, Implement, Evaluation) models (Bichelmeyer & Horvitz, 2006) as well as the role of 'Logic Models' in the field of comprehensive performance evaluation (Molenda & Russell, 2006) and their possible contributions for effective professional development evaluation was discussed. 'Realist evaluation' literature established the approach to view a program based on theory and determined the focus on contextual mechanisms that generate outcome patterns as well as causal explanations for these mechanisms and outcome patterns (Maxwell, 2004, 2012; Pawson, 2006, 2013; Pawson & Tilley, 1997).

The literature regarding effective evaluation in educational organizations further specified the focus of this study on effective evaluation of professional development and its necessary elements (Guskey, 2000). The results and implications of research studies addressing the impact of professional development on student learning were presented.

Professional development design literature, building on the seminal study of Beverly Showers, Bruce Joyce, and Barrie Bennett (1987) that analyzed thirty years of research and set the foundation for professional development activities and research afterwards was synthesized to explain the role and significance of instructional coaching. Additionally, the literature specifically pertaining to instructional coaching was reviewed to set the theoretical context for the study's underlying particular professional development design that was evaluated and to demonstrate the relevance and reason for instructional coaching (Knight, 2007). Particular characteristics of instructional coaching professional development evaluations were considered.

**The Organization as a System**

Organizational theory was unique in that it drew from many different disciplines and their perspectives. Hatch & Cunliffe (2013) summarized the multiple influences and contributions over a time span of more than two hundred years. Adam Smith in his seminal work "An Inquiry into the Nature and Causes of the Wealth of Nations" was known as the first to develop an organizational theory that became the foundation of economics, as we know it (Smith, 1776/1993). Maximum efficiency through specialization and the division of labor was its main assumption. Max Weber's work investigated the structure of authority in an organization and men's ability to rationalize within the social order (Hatch & Cunliffe, 2013; Kieser & Ebers, 2014). Formal rationalization on an institutional level lead to clear structures and competencies of the individual agent holding a position that was so clearly defined that humans became interchangeable and the organization still functioned (Hatch & Cunliffe, 2013; Kieser &

Ebers, 2014). Ludwig von Bertalanffy examined systems in different sciences and sought laws and principles of organization in the manifestation of natural systems (Hatch & Cunliffe, 2013; Laszlo & Krippner, 1998). His idea of a general system theory was first presented at the University of Chicago in 1937 and by the 1960's systems thinking was recognized as a paradigmatic effort to formulate theory across the disciplines (Laszlo & Krippner, 1998). The economist Kenneth Boulding collaborated with Bertalanffy and approached systems theory from the direction of economics and the social sciences. The strength of general systems theory was in its broad application and powerful conceptual approach to explain the interrelations of human beings. The concept of systems could be defined as a complex make up of interactive components together with the relationships among them that act within a boundary-maintaining entity or process (Laszlo & Krippner, 1998). More specific definitions required the following properties of a system: Each part of the system had an effect on the functioning of the whole, was affected by at least one other part in the system, and all possible subgroups of parts also had the first two properties (Ackoff cited in Laszlo & Krippner, 1998).

Relevant for this study was the systems approach as a process of inquiry, which differed from the traditional scientific process of inquiry by purposefully considering the embedded context at the beginning of the inquiry. It was followed by the description of the sub-wholes within the system, and then the attention was focused on the specialized parts with an emphasis to understand the structures, their compositions, and modes of operations (which was similar to the scientific method of analysis). The final step refocused again on the context, and gained knowledge was integrated to understand the

overall phenomenon, including its internal and external context (Laszlo & Krippner, 1998).

Peter Senge re-introduced systems thinking in modern management as the fifth discipline to build learning organizations (Senge, 2006). He referred to it as the discipline that integrated the disciplines and ultimately facilitated the shift from an isolated, individual perspective, to a collaborative, collective view. Schein referred to systems thinking when he addressed the connection between a learning culture and a learning leader pointing out that complexity and interdependency required a systemic analysis of the world (Schein 2010). Complex mental models took the place of simple, linear causal logic in order to learn (Schein, 2010). In the learning organization, people were continually discovering how they could create their reality and how they could change it (Senge, 2006).

In summary, the conceptual frame for this study was a systemic view of the organization with special attention to the embedded context in which all actions took place and the necessity to think systemically to collectively learn and improve practice within the organization (Laszlo & Krippner, 1998; Schein, 2010; Senge 2006).

**Performance Improvement Using a Systemic Approach**

The purpose of Human Performance Technology was to improve human performance (Pershing, 2006). The underlying assumption was that human performance was lawful in the sense that it could be engineered and systematically improved (Birchelmeyer & Horvitz, 2006). Pershing (2006) suggested that Human Performance Technology was the study of productivity improvement of organizations understood as a

systematic inquiry. Human Performance Technology focused on valuable, measurable results and considered the larger system context of people's performance (Addison & Haig, 2006). The organization was viewed as a complete system comprised of the following components: Inputs, Processes, Results and Receivers. The inputs were the resources, manpower, tools and strategies that initiated and directed actions and processes. The processes entailed the operational steps and actions to produce the desired results. Results were the outcomes of actions and processes. The receivers were the beneficiaries of the results. There were two different types of feedback, performance feedback and value feedback. Performance feedback was the feedback from within the system regarding performance and outputs while value feedback was feedback from outside of the system (Addison & Haig, 2006). All components were dependent on the conditions within the organization and its cultural setting (see Figure 2).

Figure 2. Systems Model

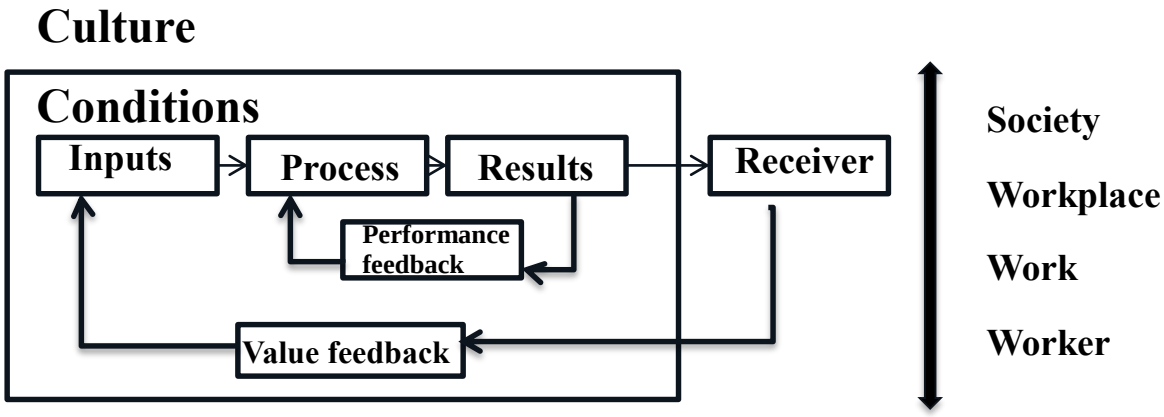



*Figure 2*. Generic systems model from "The performance architect's essential guide to the performance technology landscape" by R. M. Addison, and C. Haig, 2006, *Handbook of Human Performance Technology: Principles, Practices, and Potential* (3rd ed.), p.40. Copyright 2006 by John Wiley & Sons.

This model viewed the organization as a whole and analyzed its different components while providing a systemic framework to organize any human performance

as a process that started with inputs and lead to a performance output received by stakeholders (Addison & Haig, 2006). The evaluation of the performance was critical to measuring the level of improvement.

**Systems Approach Models for Instructional Design**

The general view of Human Performance Technology included all types of human performance. This study however focused on an instructional intervention in education (instructional coaching) and therefore this chapter reviewed literature that focused on the significance of instruction and the instructional design models that could be utilized to systemically place instruction into the larger context of performance improvement. The Human Performance Technology perspective on instruction and performance improvement was that instruction alone rarely solved performance problems because almost all performance problems were rooted in more than one cause since other factors within the system contribute to the lack of performance (Molenda & Russell, 2006).

Instructional design and development models focused on the steps to be followed when planning an intervention, from conceptualizing the problem to evaluating the effectiveness of the intervention. These models specified the necessary decisions that needed to be made and their order: They served as a procedural guide (Molenda & Russell, 2006). Gustafson and Branch (2002) identified five assumptions regarding the instructional design process and the instructional design model building when they surveyed instructional design models. By their definition there were at least five activities that instructional development consisted of: (a) analysis of the setting (b), designing a set of specifications for an effective, efficient, and relevant learner environment (c),

developing the learners' materials, and (d) implementing resulting instruction, and formative and summative evaluations of the results (Gustafson & Branch, 2002). The generic sequence was known as analyze, design, develop, implement, and evaluate and the acronym ADDIE described those types of models (Molenda & Russell, 2006). The output of each stage served as the input for the next stage (see Figure 3). One of the key attributes of this systems approach was a commitment to conduct an evaluation and revision at each step of the design and development process. Therefore Molenda and Russell (2006) pointed out that the E at the end of the acronym was to some degree a misnomer, because evaluation did not just happen at the end of the process, but throughout, and necessary adjustments were made.

Figure 3. The ADDIE Model

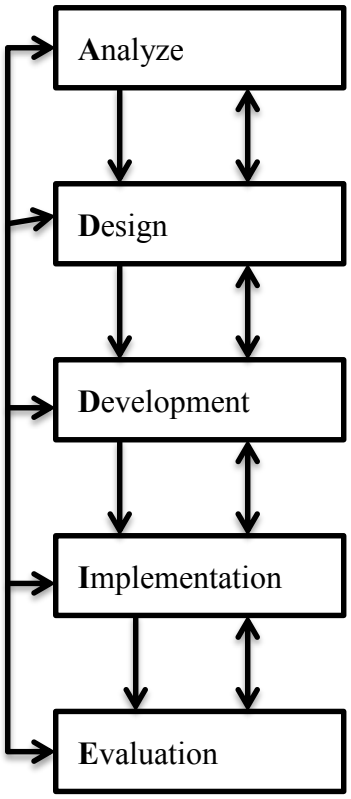



*Figure 3*. The ADDIE model. Adapted from "Instruction as an intervention" by M. Molenda, and J. D. Russell, 2006, *Handbook of Human Performance Technology: Principles, Practices, and Potential* (3rd ed.), p. 342. Copyright 2006 by John Wiley & Sons.

**The Dick and Carey Model**

According to Molenda & Russell (2006) the most widely used ADDIE type model was the one developed by Dick and Carey originally published in 1978. Their version of the ADDIE model added several levels to the analysis and development stages that specifically included the context of instruction and learning (see Figure 4).

Figure 4. The Dick and Carey Model.

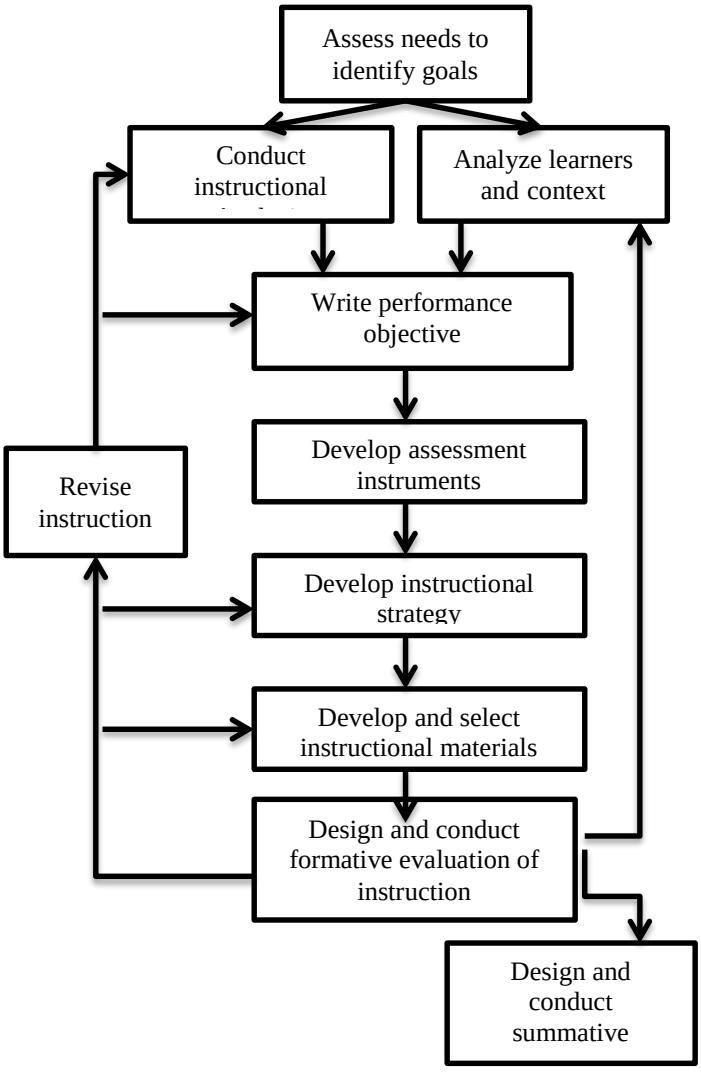



*Figure 4*. The Dick and Carey Model. Adapted from "Instruction as an intervention" by M. Molenda, and J. D. Russell, 2006, *Handbook of Human Performance Technology: Principles, Practices, and Potential* (3rd ed.), p. 344. Copyright 2006 by John Wiley & Sons.

The contribution of the Dick and Carey Model (Dick, Carey, & Carey, 2015) was its specific application for instructional interventions. It was a representation of practices in the discipline of instructional design and its purpose was to improve praxis (Dick, 1996; Dick, Carey, & Carey, 2015, Reiser, 2001). The analysis stage started out with a needs assessment to determine instructional goals and then analyzed the instructional content, the learners, the instructional context, and the context in which learning would be applied. The inclusion of context spoke to the systemic perspective of this model. The analysis stage led to the development of specific performance objectives. One of the particular strengths of the model was that after the objectives were determined it suggested a need to design the assessment instruments and methods aligned with those objectives prior to designing the actual instruction (Dick, Carey, & Carey, 2015; Molenda & Russell, 2006;). This thought of alignment between objectives of the intervention and the measure of success of the intervention in the development and planning stage was one of the key requirements the literature was stating regarding effective evaluation (Haslam, 2010; Newcomer, Hatry, & Wholey, 2010). After the assessments were developed instructional strategies were selected and the materials and procedures to implement those strategies were developed. Finally the model concluded with an evaluation and revision of the process.

**Comprehensive Performance Evaluation**

Human Performance Technology literature usually based evaluation approaches on the examination of interventions at a series of levels (Bichelmeyer & Horvitz, 2006).

One of the first was a four level approach designed by Donald Kirkpatrick in 1959. The four levels proposed were: (a) reaction, (b) learning, (c) behavior, and (d) results. The reaction level measured the satisfaction of the training participants. The learning level measured the extent to which the participants obtained knowledge, skills and attitudes as a result of the intervention. At the behavior level, the evaluator measured the extent of the change of behavior of participants based on their participation. At the results level, the value added to the organization that could be attributed to the intervention was measured (Bichelmeyer & Horvitz, 2006). Later, other level based approaches included a fifth level. Societal outcomes were first identified in Kaufmann and Keller's model (1994), a return on investment in Jack Phillips' model (1997), and organizational support in Guskey's model (2000).

The strengths of these approaches were that the levels provided a data collection and data analysis structure that allowed a way to determine the effectiveness of an intervention (Bichelmeyer & Horwitz, 2006). The levels were also easily communicated to, and understood by stakeholders. The logical progression of the levels facilitated data collection without limiting it to certain methods (Bichelmeyer & Horvitz, 2006).

However, there were certain limitations of these level based approaches. According to Bichelmeyer & Horvitz (2006), level based approaches were outcome focused, and not focused on a complex set of all elements that interact together to cause the outcomes and therefore were not helping evaluators to determine what aspects of the program were working well and to inform decisions on how to improve the program. Additionally they pointed out that these evaluation approaches had a tendency to examine how the intervention actually happened, not whether the intervention was implemented as

intended and the effect if that was not the case. Lastly, they critiqued the notion that level based approaches were not theory based regarding the theoretical foundation of Human Performance Technology as a field and suggested to base a comprehensive evaluation model on a human performance equation (Bichelmeyer & Horvitz, 2006). Bichelmeyer and Horvitz referred to the field of Program Evaluation, to look for theory-based forms of evaluations.

**The Definition and History of Program Evaluation**

A program could be defined as a set of resources and activities used and implemented to achieve common goals (Newcomer, Hatry, & Wholey, 2010). Program evaluation according to Newcomer et al. (2010) was the application of systematic methods to determine program operation outcomes. Program evaluations included ongoing monitoring of a program as well as studies of program processes or program impact. The evaluation methods used were based on social science research methodologies and professional standards. Scriven (1998) pointed out that this definition was not clarifying enough and advocated for a minimalist theory. A minimalist theory was necessary to define a field by drawing lines of demarcations or otherwise the practice suffered. According to Scriven (1998) the discipline of evaluation systematically, and objectively determined the extent to which any of three properties were attributable to the entity being evaluated: Merit, worth, or significance. Merit pertained to quality, worth to value or cost-effectiveness, and significance to importance.

Historically there were seven significant time periods in the development of program evaluation and several approaches that were used by practitioners (Hogan,

2007). Worthen, Sanders, and Fitzpatrick (in Hogan, 2007) listed several emerging trends in the field such as the expansion of the use of qualitative methods, a strong shift to combining quantitative and qualitative methods, and the introduction and development of theory based evaluations.

The difference to basic research according to Posavac (2011) was that research was concerned with questions of theoretical interest without regard to information needs of stakeholders. Theory was part of program evaluation, however, evaluation findings should meet the immediate need of decision makers and program designers. The Kellogg Foundation evaluation handbook (2004a) discussed the imbalance between program evaluation that purposed to improve the quality of the program, and program evaluation that sought to prove whether a program worked. To explain the historical context of program evaluation in the U.S. they quote Patton (2002, p.5) stating:

> Program evaluation as a distinct field of professional practice was born of two lessons….: First, the realization that there is not enough money to do all the things that need doing; and second, even if there was enough money, it takes more than money to solve complex human and social problems. As not everything can be done, there must be a basis for deciding which things are worth doing. Enter evaluation.

Another factor that lead to this imbalance was the influence of the scientific method on human services research (W. Kellogg Foundation, 2004a; Eisenhart, 2005). The Kellogg Foundation (p.7) concludes their underlying view on program evaluation:

> When the balance is shifted too far to a focus on measuring statistically significant changes in quantifiable outcomes, we miss important parts of the

> picture. This ultimately hinders our ability to understand the richness and complexity of contemporary human-services programs - especially the system change and reform and comprehensive community initiatives, which many of you are attempting to implement.

An evaluation framework then had to lend itself to meet stakeholder needs as well as consider contextual data (Kellogg Foundation, 2004a). The systemic view combined with the limitations of level based approaches lead to the development of a theory-based evaluation.

**Theory-Based Program Evaluation**

Theory-based evaluation addressed the underlying theoretical assumptions of how a program was intended to work (the program theory) and then used this theory to guide the evaluation (Rogers, Petrosino, Huebner, & Hacsi, 2000). Furthermore, it sequenced the multiple steps of a program to follow the causal path from inputs to desired outcomes and evaluated each step separately. The underlying theory provided the assumptions for causal connections between activities and particular outcomes (Bichelmeyer, & Horvitz, 2006; Rogers et al., 2000). This theory based sequencing promised better evidence for causal attribution (Rogers et al., 2000).

Stame (2004) listed various approaches of theory-oriented evaluations that differed in the assumptions about underlying theories of a program in the context of the 'black box', the space between the actual input and the desired outcomes of a program. Chen and Rossi believed that the 'black box' was an empty box, the program had no theory and its evaluation was at best a form of social accounting describing the program

and count outcomes. The purpose of a theory-based evaluation was to provide a program's missing theory (Chen & Rossi in: Stame, 2004). In contrast, Weiss (in: Stame, 2004) viewed the 'black box' as being full of theories and often more than one for the same program with an inherent confusion that had to be navigated. These theories of change had two components: (a) implementation theory, and (b) programmatic theory. The latter was based on the mechanisms that made things happen and these mechanisms were what theory-based program evaluation had to make clear, break them down in subsequent mechanisms, and test them (Stame, 2004).

**'Realist Evaluation' Approach**

The third approach Stame (2004) discussed was Pawson and Tilley's 'realist evaluation' (Pawson & Tilley, 1997, 2004). Pawson and Tilley (2004) pointed out that the commonality between theirs and other theory-driven evaluations was the view of programs as products of human imagination and a perceived course to change and improve. Evaluation then had to test that perceived underlying program theory. The core question the 'realist evaluation' asked was not 'What works?' or 'Does this program work?' but "What works for whom in what circumstances and in what respects, and how?" (Pawson & Tilley, 2004, p. 1). They regarded programs as theories, embedded in social systems and active in the sense that they required active engagement of individuals. Programs from their perspective were part of an open system and could not be kept fully isolated (Pawson & Tilley, 2004).

Realists viewed programs as sophisticated social interactions within a complex social reality. Realist evaluation explained the program through four key linked concepts:

Mechanism, context, outcome pattern, and context-mechanism-outcome pattern configuration (Pawson,2006, 2013; Pawson & Tilley, 1997, 2004). The premise of their work showed that programs didn't make things change but people did. The black box in this approach was not empty but filled with people. Independent of the approach evaluators used Logic Models to model the connections and sequence between inputs, activities, outputs, outcomes and impacts (Bichelmeyer, & Horvitz, 2006).

**Logic Models for Performance Evaluation**

Theory-based evaluation required the development of a program model that sequenced the causal path between inputs and intended outcomes (Bichelmeyer & Horvitz, 2006; Rogers et al., 2000). Logic Models were used to link short-term and long-term outcomes with the activities and processes following the underlying theoretical assumptions of the program (Kellogg Foundation, 2004b). Figure 5 shows a generic Logic Model.

Figure 5. Generic Logic Model.

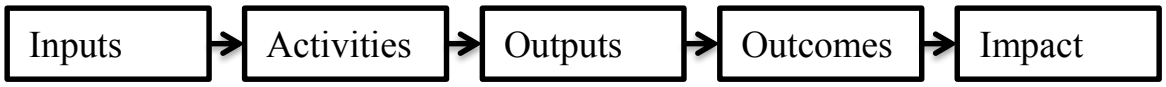



*Figure 5*. Generic Logic Model. Adapted from "*Logic Model Development Guide*" by W.K. Kellogg Foundation, 2004, p. 1. Copyright 2004 by W.K. Kellogg Foundation.

There are three prominent categories of Logic Models: Theory approach models (conceptual), outcome approach models, and activities approach models (applied) (Kellogg, 2004b). The theory approach model focused on the theory of change that underlies the program and linked theoretical ideas to explain underlying program assumptions. The outcome approach model connected specific program activities and

their outcomes. Assumed causal linkages were displayed. The evaluation needed to test the logic model (McLaughlin & Jordan, 2010; Patton, 2008) to develop the theory in practice. The use of Logic Models pointed to evaluation issues and defined key performance measurement points by its sequence. It helped with program design or improvement by closely identifying and evaluating activities. Additionally it was an excellent communication tool to stakeholders and built a common understanding of the program. Within theory driven evaluation, the Logic Model was a tool to describe the program theory and inform the evaluation (Cooksy, Gill, & Kelly, 2001; McLaughlin & Jordan, 1999, 2006). Cooksy, Gill, and Kelly (2001) noted that Logic Models were useful integrative frameworks for evaluations. They used an example that illustrated its value in facilitating triangulation and pattern matching when analyzing data. Logic Models assisted in identifying program elements for which data was collected and triangulated from different methods and sources.

**Evaluation of Professional Development: Criteria and Necessary Elements**

Guskey (2000) listed four reasons why interest in evaluation of professional development has grown: First, there had been a shift to professional development as a series of classroom-embedded learning experiences. Second, professional development had become a systematic effort for school improvement. Third, there was a need for better data to inform future professional development and educational programs. Fourth, there was an increased pressure of accountability. This pressure to demonstrate the effectiveness of educational programs was higher than ever (Gulamhussein, 2013).

Guskey (2000) defined evaluation as "the systematic investigation of merit or worth" (p. 41). It should therefore be driven by results or goals. The Joint Committee on Standards for Educational Evaluation compiled a list of 30 evaluation standards that were grouped in four groups corresponding to the four attributes of sound and fair program evaluations in education: utility, feasibility, propriety, and accuracy (Joint Committee on Standards for Educational Evaluation, 2011). The purpose of the standards was to improve the quality of evaluations. The accuracy standards clarified the bigger context in which evaluation of professional development took place. Guskey (2000) pointed out three major criticisms of evaluations of professional development. First, they were not evaluations but mere documentations, second, they were too shallow, and third, they were too brief. Accuracy standards, A2, A3, A4, and A6 (see Table 1) address those same concerns (Joint Committee on Standards for Educational Evaluation, 2011).

Table 1
*Accuracy Standards*

| | |
|---|---|
| A2: Justified Conclusions and Decisions | Evaluation information should serve the intended purposes and support valid interpretations. |
| A3: Reliable Information | Evaluation procedures should yield sufficiently dependable and consistent information for the intended uses. |
| A4: Explicit Program and Context Descriptions | Evaluations should document programs and their contexts with appropriate detail and scope for the evaluation purposes. |
| A6: Sound Designs and Analysis | Evaluations should employ technically adequate designs and analyses that are appropriate for the evaluation purposes. |

*Note*: Adapted from "*The Program Evaluation Standards* (3rd ed.)" by the Joint Committee on Standards for Educational Evaluation, 2011, p. 157. Copyright 2011 by Sage.

The standards promoted alignment between the different components of an evaluation. Intended purposes (A2), appropriate information for the intended use (A3), and the appropriate detail of context regarding the evaluation purpose (A4) were all closely connected. (Joint Committee on Standards for Educational Evaluation, 2011).

Guskey's (2000) model for evaluating professional development addressed this alignment with five levels that were hierarchically arranged from simple to more complex to ensure a higher quality of evaluation that is ultimately based on student learning outcomes. He referred to previous evaluation models and introduced Kirkpatrick's model (Kirkpatrick, 1959), as a level based approach that was developed to evaluate the effectiveness of supervisory training programs in business and industry. Guskey and Sparks (1996) introduced a model that described the factors of staff development, the content and quality of the program, and organizational climate and culture. All three factors influenced improvements in student learning outcomes. By 2002 Guskey had extended and refined the model and examined its validity through five in- depth case studies of school based professional development programs (Guskey, 2002; Guskey & Sparks 1996,). The model offered guidance to evaluations of professional development by identifying the factors that influence the relationship between professional development and student learning outcomes. Guskey adapted Kirkpatrick's evaluation model and addressed its limited use in education because of inadequate explanatory power in answering the "why" questions (Guskey, 2000). He added an additional level, organization support and change, and adapted the existing four levels to the field of professional development. His framework included the following five levels: Participants' reactions, participants' learning, organization support and change, participants' use of new

knowledge and skills, and student learning outcomes.

Each level of Guskey's framework built off the previous one and required more intricate data. Level 1 contained participants' reaction to the professional development experience. Guskey (2000) noted that it was the easiest type of information to gather and was the most common form of professional development evaluation. The questions addressed at this level were asking the participants if they felt their time was well spent, if the activities were meaningful to them and tried to capture participants' initial satisfaction with the experience. Desimone (2009) claimed that for decades studies of professional development consisted mainly of documenting teacher satisfaction, and Reeves (2010) cautioned that this focus on Level 1 feedback prevented the pursuit of higher level evaluations. Level 2 focused on measuring the knowledge and skills that participants gained. Depending on the content of the professional development activity this could be pretty simple or more complex. Measures needed to be based on learning goals of the professional development and it was necessary to develop specific criteria and indicators of successful learning prior to the activity. Pre and post assessments could capture prior knowledge and skills to measure the gains after the activity (Guskey, 2000).

Level 3 asked for organizational support and required thoughtful consideration regarding instruments to get valid feedback of the teachers whether their needs were met to undertake the difficult task of implementing their learning from level 2. Level 3 instruments required an alignment between program goals and the organizational structure in which the implementation needed to occur (Guskey, 2000).

Level 4 focused on the implementation of new knowledge and skills, the transfer of knowledge to practice. To evaluate what teachers were actually implementing the

instruments became more complex and less measurable. Guskey (2000) suggested questionnaires, structured interviews with participants, reflections from participants, portfolios, and direct observations.

Level 5 addressed the effects on students' learning outcomes, (or student achievement), the most complex level in regards to data collection (Guskey, 2000). The main question addressed at Level 5 was if and in what way the students benefitted from the professional development. Guskey (2000) stated that evaluations of professional development in education had been criticized for not providing evidence of impact on student learning. According to Guskey (2000) the specific questions that were addressed at this level were derived from the stated goals of the professional development for example higher student achievement, more positive student attitudes or perceptions or more appropriate student behaviors. He listed three major groups of student learning outcomes, namely cognitive, affective, and psychomotor skills. He suggested using student records, school records, questionnaires, portfolios, and interviews as possible data sources (Guskey, 2000).

Desimone (2009) suggested applying recent research knowledge to improve the conceptualization, measures, and methodology to study the effects of professional development on teachers and students. In accordance with Guskey and Yoon (2009), she addressed the need for better quality of studies that assessed the effectiveness of professional development to improve teachers' practice and students' achievement. She suggested a need to measure the core features of professional development, which were identified based on the research to date. The focus on critical features, according to Desimone (2009), allowed one to view the characteristics of an activity that made it

effective for teacher learning and changing practice instead of the type of activity (since there were so many complex, interactive, formal and informal opportunities for teachers to learn). She argued that there was enough empirical evidence to reach a consensus on what would be these core features. She addressed the debate regarding effectiveness, whether it meant the showing of effects on teaching practice or only when links to student achievement could be made, and argued that it was another reason that systematic features needed to be included in studies of impacts on student achievement (Desimone, 2009). She indicated that recent research reflected a consensus about some of these core characteristics of professional development.

Wilson & Berne (1999) examined research on contemporary professional development and selected existing research by three principles: it had to be high quality examples of professional development, and nominations of thoughtful work. Each nomination had to consist of a project that was committed to conducting research. The selected projects had to address the curriculum and the pedagogy, the what, and the how of teacher learning. Lastly they acknowledged that professional teaching knowledge might include knowledge of subject matter, individual students, of learning, and of pedagogy (Wilson & Berne, 1999). Their conclusions from the literature noted several themes in contemporary development and research on teacher learning: The projects involved communities of learners, teacher learning was not bound and delivered but activated, and the third commonality was the privileging of teachers' interactions with one another. Methodologically the projects' research was qualitative and labor intensive. Additionally the research projects included a connection between teacher learning and student achievement, which was something the authors noted as critical and needed.

Desimone (2009) summarized the critical features, emerging of her studies and others, were content focus, active learning, coherence, duration, and collective participation. She referred to Penuel, Fishman, Yamaguchi & Galagher's study (2007) to argue, that more recent empirical studies included these features as critical components of effective professional development. Penuel et al., (2007) used a sample of 454 teachers, who were engaged in an inquiry science program, to examine the effects of different characteristics of professional development on teachers' knowledge and their ability to implement the program. They used multiple sources of data and analyzed the data within a linear modeling framework. They were able to do this because it was a common program with a well-articulated model of implementation with fidelity and a readily available objective measure of implementation. In their methodology and research design they addressed the features Desimone (2009) mentioned. She concluded that even though the evidence of the literature supported these five core features there was no set of core characteristics that researchers regularly measured in empirical studies of professional development. She proposed a basic model that would serve as a conceptual framework that she recommended for all empirical causal studies to use. Her core theory of action for professional development followed four steps: teachers experienced effective professional development (which included the aforementioned core features). This would increase teachers' knowledge and skills and their attitudes and beliefs. Teachers would use their knowledge and skills and implement it in their practice, and those instructional changes would positively affect student learning (Desimone, 2009).

There were a number of studies that examined the relationship between features of professional development and self-reported change in teachers' knowledge and practice

(Desimone, Porter, Garet, Yoon, & Birman, 2002; Garet, Porter, Desimone, Birman, & Yoon, 2001). The connection however between professional development and student learning was complex (Guskey, 2002).

The study "Reviewing the Evidence on How Teacher Professional Development Affects Student Achievement," reviewed more than 1,300 studies identified as potentially addressing the effect of teacher professional development on student achievement in three key content areas: reading, mathematics, and science (Yoon, Duncan, Lee, Scarloss, & Shapley, 2007). Out of those studies only nine met the 'What Works Clearing House' standards further establishing the complexity of this link. In a more recent study O'Connell (2009) tried to develop a model to causally connect professional development and student achievement and had to conclude that there was no significant difference between gained scores for students between experimental group and control group.

Guskey and Sparks' model (Guskey, 2000) tried to capture the complexity of the relationship and concluded that while complex the relationship was not random or chaotic. The factors that affected this relationship could be identified and there was a clear influence (Guskey, 2000).

Several recent studies researched the impact of literacy coaching on student learning. Marsh et al. (2008) showed that literacy coaching had little effect on student learning while Garet et al. (2008) showed no effect. Biancarosa, Bryk and Dexter (2010) however showed significant effects. In their own study the coaching program was more extensive than in the previous studies. Teachers participated in a 40 hours course led by the coach and then ongoing courses of 10-12 hours of professional development throughout the year. Additionally the coaches worked one on one with teachers in their

classrooms. Their study was designed to specifically isolate the effects of a particular literacy professional development labeled Literacy Collaborate. They utilized a value added model to determine the effect of the professional development as the difference between observed and expected outcomes which was calculated by a trajectory that described an expected achievement growth for each child (Biancarosa, Bryk & Dexter, 2010). Their results showed significant gains in students' literacy learning. Students in the first year of implementation made 16% larger learning gains than observed during the baseline no-treatment period, 28% in the second year, and 32% more by the third year (Biancarosa et al., 2010). These results however were in contrast to the two previous studies (Garet et al. 2008; Marsh et al., 2008)

They listed as possible explanations the more extensive training and a more specified literacy instructional system that the coaching was based on. In conclusion they stated that evidence regarding the effects of coaching on student achievement was still slim and it was unclear if the effects of Literacy Coaching shown by their study could be observed in other coaching models (Biancarosa et al., 2010).

An evaluation framework was needed that considered Guskey's levels of evaluation with a central focus on the connection between teachers' learning, change of practice, and ultimately student achievement (Guskey & Yoon, 2009). Guskey & Yoon (2009) emphasized that the results of Yoon et al. (2007) demonstrated how little was known about this essential relationship, largely due to the lack of evidence that could be considered sound, trustworthy, and scientifically valid. Guskey and Yoon (2009) specifically state however, that the results of the study did not indicate that alternative professional development designs such as coaching did not work but that "rather, the

results illustrate at this time, we simply have no reliable, valid, scientifically defensible data to show these strategies do work" (p. 498). Their true value had yet to be determined. They charge the field of professional development to critically assess and evaluate the effectiveness of what they do and point out the necessary alignment between the specific goals of the professional development, the selection of appropriate evidence, and the collection of meaningful and scientifically defensible data. Furthermore they requested replicable measures of new strategies and practices and stated that the implementation of any new strategy should begin with carefully designed pilot studies designed to test their effectiveness. Lastly they called for greater rigor in the study of professional development to improve precision of studies of the relationship between the professional development, the effect on teacher practices, and the effect on student learning outcomes (Guskey & Yoon, 2009).

**Context and Purpose of Professional development Design**

According to Showers, Joyce, and Bennett (1987) the purpose of professional development design was "to create the conditions under which sufficient levels of knowledge and skill were developed to sustain practice and to provide the conditions that support practice until executive control has been achieved and transfer has occurred" (p.84). The highlights of the meta-analysis of nearly 200 research studies showed that almost all teachers benefited in their classroom practice when training includes four parts: (a) presentation of theory, (b) demonstration of new strategies, (c) initial practice, and (d) prompt feedback. Teachers were more likely to implement and use new strategies and

concepts if they received coaching; either expert or peer (Showers, Joyce, & Bennett, 1987).

The larger context of professional development design however was to implement sustainable change (Hawley & Valli, 1999; Reeves, 2010). If everything had been well there would be no need to change practice. The need for professional development was usually triggered by an identified area in which change needed to occur to improve student achievement, school culture, attendance, self-efficacy, school climate or any other area. In school districts, the design of professional development often followed particular overarching district goals, which in return were driven by initiatives such as the secondary school reform (The Connecticut Plan - School Reform Draft, 2007) for districts in Connecticut. The Connecticut Plan "creates expectations for pre-service teacher training and professional development of experienced teachers" (p. 23). The plan required a close collaboration between the state institutions, the public schools, and the business community to ensure that teachers were prepared in content areas and instructional practices to meet the requirements of the proposal. Professional development design therefore did not occur in isolation but within a bigger context (The Connecticut Plan - School Reform Draft, 2007)

To facilitate the sustainable changes necessary to meet requirements determined by the larger context in which teaching and learning takes place, change theory suggested that it was a slow process and required long-term change models (Fogarty & Pete, 2007; Guskey, 2002). In analyzing the large scale reforms in Seattle, Milwaukee and Chicago and their failure to change practice on a large scale Fullan (2006) attributed the failure to a lack of focus on what needed to change in instructional practice and what it would take

to bring about these changes in classrooms across the districts. In his theory of action, Fullan (2006) listed seven premises that underpinned the use of change knowledge, amongst them capacity building with a focus on results and, learning in context. Professional development then needed to be designed to promote capacity building and facilitate learning in the context in which teachers worked to ensure sustainable change in teacher practice. He stated a call to radically change the concept of what teacher learning should entail based on five ideas (Fullan, 2007): Professional development as a term was an obstacle and had run its course, a focus on Elmore's (2004) conclusion that improvement required "learning to do the right things in the setting where you work" (p.73), student learning depended on continual teacher learning, the necessity of teacher collaboration, and the role of teacher's working context and conditions to facilitate teacher learning. Traditional workshops, courses, and programs as external ideas could not result in change (Loucks-Horsley, Hewson, Love, & Stiles, 1998). "These activities are not useless, but they can never be powerful enough, specific enough to alter the culture of the classroom and school" (Elmore, 2004; p. 35). Cole (2004) viewed traditional professional development as a great way to avoid change. He based the ineffectiveness of most professional development to improve teacher practice on the observation that the professional development was usually an externally provided training program, most of these programs were not designed to produce observable teacher practice change, and generally only a few teachers per school attended these training programs, and therefore had a limited capacity to bring school wide change. Wei, Darling-Hammond, Andree, Richardson, & Orphanos (2009) examined nationally representative data from the National Center for Education Statistics' 2003-04 Schools

and Staffing Survey and reported that nearly half of all U.S. teachers were dissatisfied with their opportunities for professional development and noted, that the type of professional development model influenced the likelihood to effectively facilitate change. They indicated that effective professional development models supported a new paradigm of teacher professional learning (Desimone et al., 2002; Garet et al., 2001; Wei et al., 2009).

Guskey (2002) mentioned two main factors that professional development programs overlooked and therefore were ineffective: The reason for teachers to be motivated to engage in professional development, and the process by which change in teachers typically occurred. Earlier studies presumed that to change teacher practice one had to first change attitudes and beliefs about certain teaching practices through professional development and then the teachers would change their practice. Guskey's (2002) alternative model assumed that professional development changed teacher practice, which would affect student outcomes. Better student outcomes would affect teachers' change of their attitude and beliefs about their teaching practices. It was then not the professional development that changed the practice but the fact that the teachers had seen that the new practices worked, through improved student outcomes (Guskey, 2002).

Historically, different professional development models evolved from one-day presentations, to district wide professional development plans (multiple workshop days over time), site based professional development (training on the building level), communities of learners (involvement of collaborative teams to facilitate professional learning), and individualized professional learning plans (Fogarty and Pete, 2007). In

most districts the professional development contained a majority of the models simultaneously.

**Acquisition of Knowledge and Skills**

According to Joyce and Showers (2002) the focus of professional development was the acquisition of knowledge and skills and the achievement of transfer. In the professional development context, the focus was on the learning experiences of adults. When designing professional development and choosing a particular venue, program, or activity the characteristics of adult learners should have been considered. In their professional development and revised version of Malcolm Knowles' seminal work *"The Adult Learner: A Neglected Species"* (1973) Holton and Swanson incorporated advancements on the core principles of Knowles' pioneering theory of andragogy (Knowles, Holton, & Swanson, 2005). They listed several assumptions about adult learners: They needed to know the learners' self-concept, the role of the learners' experiences, readiness to learn, orientation to learning, and motivation. Adults needed to know why they should learn something. Professional development had to ensure that the purpose for the learning was clear to teachers and that they perceived it as beneficial. The experiences adult learners brought to the teaching and learning setting were both enriching but also potentially limiting because of formed habits, established mindsets and bias. Professional development needed to tap into prior experiences but also helped adult learners to identify their habits and biases and be open to new approaches. Adult learners were life-centered, task-centered or problem-centered learners which meant they would learn what was relevant to their life context. For teachers, this meant that professional

development should address relevant content for practice for the teacher's school and classroom context. These conclusions coincided with Guskey's (2002) account that teachers valued professional development because they believed that it would expand their knowledge and skills, contribute to their professional growth, and allow them to be more effective with their students. On the other hand he pointed out that teachers tended to be pragmatic and their focus was on specific, concrete, and practical ideas that directly related to the day-to-day operation of their classrooms. Professional development that didn't address those needs was unlikely to succeed (Guskey, 2002).

**Achievement of Transfer**

As previously mentioned, research identified certain elements of professional development designs that facilitated sustainable change in practice (Cole 2004; Desimone et al., 2002; Elmore, 2004; Fullan, 2006; Garet et al., 2001). The re-occurring themes were the provision of contextual learning in the setting in which teachers were actually working and the provided timeframe within which teacher learning was supported.

Joyce and Showers (2002) estimated that for a complex model of teaching about twenty-five teaching episodes, during which the new strategy was used, were necessary for transfer to be achieved and to not be lost through disuse. Their research (Joyce & Showers, 2002) showed that the coaching process enabled nearly all teachers to sustain the practice and to gain executive control over a large range of instructional practices. Based on their research and experience they connected training components and training outcomes and concluded that coaching would increase transfer (executive implementation) from 5% (after study of theory, demonstrations, and practice) to 95%

(Joyce & Showers, 2002). Coaching according to their early studies contributed to transfer in five ways: more frequent practice, more appropriate use of newly learned strategies, exhibition of greater long term retention of knowledge about and skill with strategies, greater likelihood to explain new models of teaching to their students, and clearer cognitions regarding the purposes and uses of new strategies.

Fogarty and Pete (2007) listed seven strategies (Fogarty & Pete, 2007; Perkins, & Solomon, 1987) that promoted transfer from staffroom training to classroom practice to consider implementing when "coaching for transfer" (p. 118): Learning about transfer theory, setting expectations for transfer, modeling with authentic artifacts, reflecting on levels of transfer, plotting applications with the tiny transfer book, try something immediately, and having a dialogue with hugging and bridging questions. They refer to two types of transfer (Perkins & Solomon, 1987): (a) simple transfer, and (b) complex transfer. The difference lied in the level of closeness of the new learning compared to the learning in the original situation. The more students practiced for example, the more automatic becomes the transfer of the information. Complex transfer however was not naturally occurring and required thinking about the application of the learned skill in the new context as well as bridging (Fogarty & Pete, 2007). For professional development design this they applied this to the workshop scenario where teachers were introduced to a new instructional strategy that they were then expected to regularly use in their classroom where practicing regarding that strategy had not occurred. However, if teachers could learn the new instructional strategy in their own classroom setting with the support of a coach and practice it throughout the school year with multiple coaching sessions to support the learning, the complex transfer could occur.

**Impact on Student Learning Outcomes**

The targeted content of professional development experienced a shift from self-reported teacher needs to a more student centered perspective due to the value that was placed on student achievement data and its role to inform instructional practices and therefore professional development (DuFour, 2010; Reeves, 2010; Tallerico, 2005). Teaching and learning theories that focused on student centeredness had received increased attention (Beane, 2002), and the educational political environment emphasized students' learning outcomes and the thought that students' learning needs should determine professional development activities (Killion, 2002). Tallerico (2005) discussed two approaches to determine a more student-centered content of professional development: One was strongly curriculum based (micro level perspective) and one was focused on data based decision making (macro level perspective).

Professional development needed to connect adult learning with student needs. Tallerico (2005) listed a number of professional development activities which focused on everyday curriculum that students struggled with most: teachers who shared the same content would discuss everyday lessons based on formative assessments and identification of necessary modifications, creating and administering common assessments, and as a third source of adult learning content, the actual student work. Collaborative problem solving and action research would be applicable design options for professional development focused to increase teachers' abilities to conduct these activities.

From a macro level perspective data informed the content choice of professional development. Report card grades, district wide assessment grades, dropout rates,

graduation rates, attendance and discipline data, standardized test scores, career interests and student success plans. She noted that all were data sources to identify areas of need (Tallerico, 2005).

Research had firmly established the influence of teachers on student learning, as well as the necessity for professional development to facilitate change that resulted in improved student achievement (Hawley & Valli, 1999). Reeves (2010) referred in his introduction to Guskey's (2000) landmark call to move from evaluations based on participants' reactions to evaluations based on student learning. Pfeffer and Sutton's (2000) work summarized research findings by stating "the most salient variable in improving student achievement was not the brand name of the program but the degree of implementation of the program. In brief, it was practices and people, not programs that made the difference for student achievement" (p.3).

Joyce and Showers (2002) listed four conditions that had to be present to significantly affect student learning: (a) a collaborative community of professionals; (b) curricular and instructional strategies which were probable to affect student learning as the main content of professional development; (c) significant changes in what was taught, how it was taught, and the social climate; and (d), processes that enabled educators to implement their learning (p.4).

Showers and Joyce emphasized professional development that influenced the "what" and the "how" of teaching with consideration of the social climate of the school due to the proximity of the school environment to the student and its impact on the development of students. Professional development could and should have had an effect on student learning. Student learning according to Joyce and Showers (2002) could be

measured in many ways and they listed examples such as course grades, attendance, measures of conduct such as numbers of referrals to special education, suspensions, curriculum-relevant tests, and standardized tests as well as the analysis of student work and portfolios. The exemplary cases they introduced all demonstrated that student learning could be affected by professional development but they also concluded that most of the professional development that was offered was probably not affecting it.

Several later research studies were able to show that coaching could have an effect on student achievement. In 2006, The Learning Network reported findings from Battle Creek, Michigan and showed steady growth in student achievement over five years in 4$^{th}$ grade students' reading scores from 29% achieving proficiency in 1999 to 86% proficiency in state standardized exams (Cornett & Knight, 2009; Killion et al., 2012). The Learning Network training program included teacher leaders who acted as literacy coaches. However, in the research design there was no control group present. Cornett and Knight (2009) and Killion et al. (2012) noted that due to the methodologies used they could not be sure that the teacher leader coaches caused change.

There were a number of later studies that sought to study the impact of professional development that included some element of coaching with rigorous randomized trial studies. Allen, Pianta, Gregory, Mikami, and Lun (2011) conducted a randomized control trial of a coaching program, the My Teaching Partner-Secondary program (MTP-s). The program focused on teacher-student interactions to enhance student motivation and achievement. The coaching program contained initial workshop based training, one year of personalized web based coaching (teachers submitted videos that coaches reviewed) and was followed up with a brief workshop. The study included

2237 students, and 78 teachers from 12 schools who participated for 13 months in MTP-s and for two years in the evaluation of the program. Results indicated a non-significant effect on end-of-year test scores but a positive significant effect in the post-intervention year. The net gain to the control group equated to an average increase in student achievement from the $50^{th}$ to the $59^{th}$ percentile for a student moved from the control group to the intervention group (Allen et al., 2011).

Garet et al. (2010, 2011) conducted The Middle School Mathematics Professional Development Impact Study in 12 districts. Each district provided 4 to 8 schools, 77 schools in total, with 195 teachers and 11,479 students. The study measured teacher knowledge of rational numbers content pedagogy, teachers' instructional practices, and student achievement in rational numbers. The professional development included summer workshops, 1-day seminars during the school year and coaching visits. The results of this randomized experimental design showed after one year of implementation that the program had no statistically significant impact on teacher knowledge, a statistically significant positive impact on the frequency with which teachers engaged in activities that elicited student thinking, and did not have a statistically significant impact on student achievement (Garet et al., 2010). They stated the question whether or not the study's outcome measures captured aspects of teacher knowledge and practice that could be associated with student achievement and demonstrated again how difficult it was to account for that impact. After the second year of implementation the results were the same, no significant impact on teacher knowledge or student achievement (Garet et al., 2011).

These large-scale experimental studies all met the requirements of effective professional development design: content focus, active learning, and duration (Desimone, 2009). They sought to measure its impact on the essential components of teacher knowledge, teacher practice, and student achievement. The followed a randomized control trial that allowed causal inferences to be made about the effectiveness of the professional development strategies. The contrasting results showed that in some cases the professional development had significant effects and in others it did not. The limitations and conclusions of the studies further emphasized the difficulty to make causal connections between professional development and student achievement (Allen et al., 2011; Biancarosa et al., 2010; Garet et al., 2010, Garet et al., 2011).

**Instructional Coaching**

After interviewing more than a 150 teachers across the United States, Knight (2007) concluded that the issue wasn't that teachers resisted change so much but that they resisted poorly designed change initiatives. This reinforced the notion mentioned earlier, that traditional professional development was not effective (Elmore, 2004; Reeves 2010; Joyce & Showers, 2002). Knight (2007) cited Bush whose study showed a maximum of 10% implementation rate of traditional professional development. In Knights experience more than 90% of teachers embraced and implemented programs, that improved students' experiences in the classroom, when they received an appropriate amount of support (2007). His research had grown into the promotion of an approach referred to as instructional coaching. According to Knight (2007) instructional coaches worked with teachers to help them incorporate research-based instructional practices. He listed four

issues that should be considered to obtain focus for the collaboration between coach and teacher: (a) behavior; (b) content knowledge; (c) direct instruction; and (d) formative assessment. All of these issues were core focus areas for the coach to work on with the teacher. Additionally instructional coaches needed to know a lot about teaching and how to establish emotional connections with their teachers. Knight (2007) referred to this as the Partnership Mind-set, which was built around the seven core principles: quality, choice, voice, dialogue, reflection, praxis, and reciprocity. Equality meant that collaborating teachers views were equally important to the coach's views. Choice meant that teachers had a choice in the how and what of their learning. Voice meant, that teachers expressing their opinion about learned content were encouraged. Lastly dialogue meant that coaches not only tell but also listen (Knight, 2007).

Spaulding & Smith (2012) pointed out the importance of the instructional coach working with a school community to be successful in improving classroom instruction and eventually student performance. Instructional coaching, especially because it was classroom embedded and ongoing unlike traditional professional development activities, was much more exposed to the context of the actual school and building than professional development at an offsite location. Within a school building the building administration had the role of instructional leaders (Spaulding & Smith, 2012). The responsibilities of administrators could be categorized in priority one and priority two issues. Priority one issues according to Spaulding & Smith (2012) were safety and welfare issues without which learning could not take place (school suspensions, student attendance, staffing, office referrals, building maintenance issues, and grievances). Priority two issues dealt with the main purpose and vision of the school, the academic achievement of all students.

Examples of priority two issues were implementing research-based curriculum, data analysis of student achievement data and student work to inform instruction and to monitor growth. Additionally, priority two issues included classroom observations and relevant feedback to teachers on a regular basis, creating effective classrooms and learning environment conducive to all learners, differentiated instruction, and student engagement. Instructional coaches became a support to the instructional leadership team of the school to meet the needs of teachers and students that could often not be met due to priority one issues. It required an organizational commitment to coordinate the greater vision and needs of the building with the instructional coach servicing the teachers. It was important to understand that an instructional coach was not just someone with a bag of teaching tricks but someone who could collaborate with all staff and administration in the building, could communicate effectively always keeping the overall goals, purpose, and mission in mind so that all efforts moved in a concise direction (Spaulding & Smith, 2012).

Guskey (2002) listed three principles that stemmed from his teacher change model and should be considered when designing professional development: Recognize that change was a gradual and difficult process for teachers, ensure that teachers received regular feedback on student learning progress (teacher practice changed when there was evidence of increased student learning due to it), and provide continued follow-up, support and pressure. Instructional coaching as a professional development design model was able to meet these requirements, due to being embedded in the building, the collaborative teamwork and the ongoing nature of it.

**Evaluation of Instructional Coaching**

The literature specifically regarding the evaluation of instructional coaching was sparse. Biancarosa et al. (2010) showed that literacy coaching affected teacher practice and student achievement and Allen et al. (2011) showed that a web based coaching program to improve teacher –student interactions impacted student achievement positively. Garet's et al. (2011) large-scale empirical studies failed to show a significant impact of the mathematics professional development that included coaching, on teacher practice and student achievement. These studies did not evaluate Knight's (2007) instructional coaching model.

Cornett and Knight (2009) synthesized the research to date regarding coaching and stated that out of more than 200 publications describing some form of coaching related research, most did not meet the standards of rigorous research. They emphasized that there were different types of coaching: Peer Coaching, Cognitive Coaching, Content Coaching, Literacy Coaching, and Instructional Coaching. Cornett and Knight (2009) noted that Instructional Coaching focused on providing classroom embedded support for teachers to enable them to implement scientifically proven teaching practices.

They summarized the results of their own research regarding the effect of instructional coaching. They surveyed 107 teachers in the 2003-2004 school year in the Topeka School District, who had viewed a model lesson from a coach and the results suggested that the teachers believed they benefitted from it. They conducted 13 interviews with teachers who had collaborated with a coach in middle school and in each of the interviews the teachers stated that the model lessons were an essential part of the coaching process (Cornett & Knight, 2009). In the summer of 2004, 82 teachers in

Topeka attended summer workshops and then had Instructional Coaches in their schools to provide additional support in implementing the new teaching practices. In October 2004 the coaches conducted classroom visits and reported that 70 out of the 82 teachers who received coaching were implementing those practices. 85% of the teachers implemented the new practices they had learned during the summer (Cornett & Knight, 2009).

They concluded that research regarding Instructional Coaching, particularly the studies of implementations, needed further study. Their implementation study only contained self-reported data. Most importantly, they stated, research on the impact of Instructional Coaching on student achievement was needed.

They proceeded to study the impact of Instructional Coaching with a mixed method study in which a simple between-subject experimental design was employed (Cornett & Knight, 2009). Fifty teachers participated in the study and completed it. The study focused on the Unit Organizer Routine, one of several routines of the Content Enhancement Series developed at the University of Kansas Center for Research and Learning. The participants all attended a workshop where they were taught how to use the teaching routine (one 1.5 hour session after school). After the workshop the participants were randomly assigned to one of two conditions: receive instructional coaching support or not receive no support following the workshop. The results showed that teachers who were supported by an instructional coach used the teaching routine more than teachers who only attended the workshop. They listed a number of limitations of the study and the study did not address an impact on student achievement, only on a change of teaching practice (Cornett & Knight, 2009).

Several school districts have implemented instructional coaching programs and have either developed evaluation plans or have started to evaluate the programs. The Best Practices In Evaluating the Impact of Instructional Coaching in Teaching and Learning Report (APQC, 2011) summarized the results of instructional coaching evaluations in Allen Independent School District; Allen, Texas, Dysart Unified School District; Surprise, Arizona, Iredell-Statesville Schools; Statesville, North Carolina, and Springfield Public Schools; Springfield Missouri. The districts had a coaching program that had been in place for more than two years, and had documented results that showed positive impact on teaching and learning. The study came to the following conclusions: Isolating the results of coaching was a challenging task, all districts reported significant improvements tied to their instructional coaching program, coaching programs supported by positive data had a high success rate of surviving budget cuts, and most districts used a variety of measurement tools to determine success of instructional coaching (APQC, 2011).

Friendship Charter School (2011) developed a performance coach plan to evaluate their coaches' performance mainly based on a joint supervision of the coaches by the principal and the Director of Professional development. Ithaca City School District (2013) developed an evaluation plan for their instructional coaching program. They developed an ADDIE model, a Logic Model and measurements. The measurements included teacher and student measures. Teacher measures included Danielson's evaluation framework and a set of reflection questions regarding teachers' feedback about the coaching model. They divided student performance in three different measures. The instruments included pre and post surveys, benchmark tests, pre and post observation

ratings, student tests scores, implementation checklists, and more. The General Electric Foundation evaluated the coaching initiative in Erie Public Schools (2013). Their findings were based on multiple surveys and included: evidence of professional development, preparation to implement Common Core State Standards, English, language arts and mathematics classroom practices, teacher knowledge, perceptions of coaching, and other factors influencing implementation. The results showed that there was a significant increase in professional development in Common Core State Standards and that it was well received. Student achievement was not addressed however. In summary the studies all employed multiple instruments, mostly surveys, interviews and some type of protocol to observe the frequency of implementation.

Regarding the use of program evaluation tools Danks (2011) suggested to evaluate instructional coaching using the Analysis, Design, Develop, Implement, Evaluate model (ADDIE) without providing empirical evidence though that she applied the model herself to effectively evaluate instructional coaching. The Analysis, Design, Develop, Implement, Evaluate model was a five-step instructional design model, as mentioned earlier, borrowed from the field of HPT and was commonly used to develop, implement, and evaluate performance improvement services (Danks, 2011). Some of the core values and principles that governed the use of the ADDIE Model in HPT included a focus on outcomes, a systems perspective, an intent to add value, and an effort to be systematic in all aspects of the design of interventions and solutions. According to Danks (2011) using the (ADDIE) Model of instructional design, allowed to evaluate the impact and effectiveness of instructional coaching models.

Some of the key questions about instructional coaching that remained for stakeholders regarded the proper alignment to student, teacher, principal and federal needs (Danks, 2011). Additionally stakeholders wanted to know how an instructional coach evaluated his or her effectiveness and impact and how did an instructional coach personally model the usage of a systematic learning process and data-enhanced decision making (Danks, 2011). Table 2 shows the ADDIE model for the instructional coaching professional development.

Table 2
*ADDIE Model for Instructional Coaching Professional Development - Generic*

| Process/Step | Key Activities |
|---|---|
| **Analyze**<br>Conduct needs analysis | • Collaborate with central office staff to determine district needs;<br>• Collaborate with site administration to determine perceived needs of staff<br>• Analyze school performance results and areas of improvement |
| **Design**<br>Design the support model | Draft logic model based on data analysis |
| **Develop**<br>Develop key work functions | Collaborate with central office and site stakeholders to develop the key work functions |
| **Implement**<br>Implement the key work functions | Implement the key work functions |
| **Evaluate**<br>Evaluate the impact of each key work function | • Analyze quantitative and qualitative data.<br>• Present findings to stakeholders |

*Note:* Adapted from "The Addie Model: Designing, evaluating instructional coach effectiveness," by S. Danks, 2011, *ASQ Primary and Secondary Education Brief, 4(5*), p.2. Copyright 2011 by ASQ.

The first step in the model was the Analysis (step A). The Logic Model (step D) was developed based on the analysis in step one. The Logic Model as mentioned earlier was a picture of how the program worked and a key element of a theory based evaluation approach. The benefits of using a logic model were a clearer focus on connections between intentions and outcomes, and an effective way to lay out the more measurable interim outcomes on the way to more intangible long-term outcomes/impact. The development process itself was valuable because it required collaboration with stakeholders and clarification of underlying rationale; thus gaps could be identified (Kellogg Foundation, 2004a).

In summary the literature suggested to use multiple methods and instruments, and to develop an ADDIE and Logic Model and to try to measure effects on student achievement.

**Summary**

The systems approach, as a process of inquiry, purposefully considered the embedded context at the beginning and end of the inquiry to not only understand the structures, compositions and modes of operations of an organization; but the overall phenomenon in its internal and external context (Laszlo & Krippner, 1998). The field of Human Performance Technology systematically analyzed an organization to evaluate its performance. The general model viewed the organization as a whole and analyzed its different components while providing a systemic framework to organize any human performance as a process that started with inputs and lead to a performance output received by stakeholders (Addison & Haig, 2006).

Human Performance Technology used instructional design models to analyze instructional interventions. Instructional design and development models specified the necessary decisions that needed to be made and their order: They served as a procedural guide (Molenda & Russell, 2006). It provided a systematic sequence to plan, implement and evaluate an intervention. The ADDIE models typified models that met these criteria (Molenda & Russell, 2006). Danks (2011) suggested the use of the ADDIE model to evaluate instructional coaching because it allowed the evaluation of the impact and effectiveness of instructional coaching models. For the actual evaluation of the intervention the Human Performance Technology literature referred to level based

approaches that provided a data collection and data analysis structure that allowed a way to analyze an intervention at each of the levels (Bichelmeyer & Horwitz, 2006). However, the limitations of these level based approaches, which included Guskey's (2000) model, were, that they did not focus on a complex set of all elements that interacted together to cause the outcomes and therefore did not help evaluators determine what aspects of the program were working well and to inform decisions on how to improve the program. Level based approaches were not theory based and Bichelmeyer and Horvitz referred to the field of Program Evaluation to look for theory-based forms of evaluations (2006). Theory-based evaluation addressed the underlying theoretical assumptions of how a program was intended to work (the program theory) and then used this theory to design the evaluation (Rogers, Petrosino, Huebner, & Hacsi, 2000). The underlying theory provided the assumptions for causal connections between activities and particular outcomes (Bichelmeyer, & Horvitz, 2006; Rogers et al., 2000). This theory based sequencing promised better evidence for causal attribution (Rogers et al., 2000). Logic models were used to sequence the causal path between inputs and intended outcomes (Bichelmeyer & Horvitz, 2006; Rogers et al., 2000).

The organizational theory literature, and Human Performance Technology literature suggested the use of these evaluation tools to ensure a systemic, comprehensive approach that was theory based and considered the context in which the professional development would take place. 'Realist evaluation' added the distinct viewpoint on how an intervention brought about change (Pawson & Tilley, 2004). 'Realist evaluation' (Pawson, 2013) stressed the concepts to explain a program: Mechanisms, context,

outcome pattern, and context-mechanism-outcome pattern configurations or as Maxwell (2004) stated: context + mechanism = outcome.

The professional development literature (Desimone, 2009, Guskey 2000) suggested an underlying core theory of action that included the following steps: Teachers experienced effective professional development. The intervention increased teachers' knowledge and skills as they used their knowledge, skills, and attitudes to improve the content of their instruction. The change in instruction led to increased student learning (Desimone, 2009). Core elements of effective professional development according to the literature contained; (a) presentation of theory, (b) demonstration of new strategies, (c) initial practice, and (d) prompt feedback. Teachers were more likely to implement and use new strategies and concepts if they received coaching; either expert or peer (Showers, Joyce, & Bennett, 1987). It was essential that teachers could practice their learning in the setting they actually worked in (Cole 2004; Elmore, 2004; Fullan, 2006). Joyce & Showers (2002) showed that the coaching process enabled nearly all teachers to sustain the practice and to gain executive control over a large range of instructional practices.

Guskey's (2000) five levels of evaluation coincided with Desimone's (2009) proposed core theory of action for professional development: Participants' reactions, participants' learning, organization support and change, participants' use of new knowledge and skills, and student learning outcomes. Thus, an evaluation framework was needed that considered Guskey's levels of evaluation with a central focus on the connection between teacher learning, change in practice, and the effect on student achievement. (Guskey & Yoon, 2009). Guskey & Yoon (2009) noted that little was known about this relationship, due to the lack of empirical evidence. Guskey and Yoon

(2009) specifically stated that at this time, there were no reliable, valid, scientifically defensible data to show that alternative professional development strategies such as instructional coaching worked. They charged the field of professional development to critically assess and evaluate the effectiveness of professional development and pointed out the necessary alignment between the specific goals of the professional development, the selection of appropriate evidence, and the collection of meaningful and scientifically defensible data. Furthermore they requested replicable measures of new strategies and practices. They called for greater rigor in the study of professional development to improve precision of studies of the relationship between the professional development, the effect on teacher practices, and the effect on student learning outcomes (Guskey & Yoon, 2009).

Although much work has been done to develop a valid evaluation framework that included key features of effective professional development and addressed and sought to measure the impact on teacher practice and student learning outcomes (Allen et al., 2011; Desimone, 2009; Garet et. al, 2008, 2010, 2011) a comprehensive framework to evaluate alternative professional development designs such as instructional coaching, did not exist (Guskey & Yoon, 2009). Further, research was therefore needed to meet these requirements of effective professional development evaluation, including alignment of goals and collected evidence in a meaningful and scientifically defensible way, to evaluate contextual, alternative professional development designs.

## Chapter 3: METHODOLOGY STATEMENT & METHODS/PROCEDURES

The field of evaluation has defined mixed methods designs early on as "those that include at least one quantitative method (designed to collect numbers) and one qualitative method (designed to collect words), where neither type of method is inherently linked to any particular inquiry paradigm (Greene, Caracelli, & Graham, 1989; p. 256). Greene (2007) broadened this definition to a "mixed methods way of thinking' and stated (p. 20)

> "The primary purpose of a study conducted with a mixed methods way of thinking is to better understand the complexity of the social phenomena being studied. In a mixed methods way of thinking, better understanding can take various forms, more than one of which may be invoked in any given study"

This study presents a unique form of mixing beyond quantitative and qualitative collection and analysis of data strands, but by mixing on a theory level (substantive theory stance; Greene, 2007).

### Mixed Research Design: A Substantive Theory Stance

The substantive theory stance in mixed methods inquiry originated according to Greene (2007) from theory based evaluation. As outlined in chapter two theory-based evaluation emphasized the underlying theoretical assumptions of how a program was intended to work (the program theory) and then used this theory to guide the evaluation (Rogers, Petrosino, Huebner, & Hacsi, 2000). 'Realist evaluation' further combined the realist paradigm with the rationale that evaluation had to test the perceived underlying program theory (Pawson & Tilley, 1997; Pawson, 2006). The paradigmatic emphasis is

the theories, and the evaluation methods are driven by the theories of the program (Pawson & Tilley 1997). Greene (2007) noted that ”data are not analyzed and aggregated by method; rather, data analysis is framed and organized by concept or theory” (p.74).

In this study the data collection is framed by Guskey's (2000) model of professional development evaluation and Desimone's (2009) suggested core action theory of professional development. Qualitative and quantitative data is ultimately interpreted through Context-Mechanism-Outcome configurations as determined by the underlying program theory. As Greene (2007) pointed out, evaluations present results per evaluation question not for each method separately. Consequently this study's evaluation questions were based on the underlying theories and the results presented for each evaluation question.

**Causal Explanations and 'Realist Evaluation'**

This mixed methods program evaluation of professional development strategies was theory based using a 'realist' approach first proposed by Pawson and Tilley in 1997. The focus of this study was to develop an evaluation framework and to create applicable instruments to conduct a 'realist evaluation' of the effects of instructional coaching in mathematics.

The core question behind traditional program evaluation was causal in nature: Did the program work? From a practical perspective one of the greatest challenges in evaluating teacher professional development is to determine patterns for causality. The evaluation designs to empirically do so are beyond the scope of most state and local evaluation efforts (Haslam, 2008). The comprehensive review of the literature clearly

showed that demonstrating causality between professional development and successful implementation of teaching practices and student achievement is extremely complex and the studies that attempted to show this relationship often relied on poorly developed measures of the professional development features or self-report survey instruments (Eisenhart, 2005; Guskey, 2000; Guskey & Yoon, 2009; Yoon et al, 2007)

Eisenhart (2005) discussed four research designs to pursue questions regarding causation in education and agreed with other researchers (Howe, 2004; Maxwell, 2004) that good descriptive knowledge was essential if causal analysis was to succeed. To understand whether and how x causes y it was first necessary to know what x and y were and how x in actual practice could influence y. As an alternative, process views of causality based on a realist perspective acknowledge that human behavior and social interactions are often the causes of educational effects. Proponents of the realist perspective do not reject the regularity view but emphasized that "it left much out that should be included in research that focused on human activity" (Eisenhart, p. 255).

Eisenhart (2005) referred to the three kinds of research questions as stated in the Scientific Research in Education (SRE) report designed to answer: (1) What was happening, (2) was there a systemic effect, and (3) why or how was it happening?

According to Howe (2004) not only the second kind of questions were asking for causality (the viewpoint of the SRE) but all three. Eisenhart (2004) concluded that while proponents of experimental research in education deemed experiments the gold standard to answer causal questions, "qualitative and other research designs and methods were needed to answer the harder causal questions about the antecedents that lead to effects and about the process that makes antecedents work" (p.255).

Causality within the realist approach referred to causal mechanisms and processes that were involved in particular events. The mechanism was responsible for the relationship, not a variable but an account of the relationship, which opened the door for causation in qualitative research (Maxwell, 2012; Maxwell & Mittapalli, 2010).

In 'realist evaluations' programs are complex social systems and the results should answer the question 'what worked for whom in what circumstances and in what respects' (Pawson, 2013; Pawson & Tilley, 1997, 2004). The focus is on the connection between mechanisms, context, and outcome patterns. Mechanisms (M) describe "what it is about programmes and interventions that bring about any effects" (Pawson & Tilley, 2004: p. 5). The mechanisms are the engine to explain occurrences and what actually happens. Context (C) "describes those features of the conditions in which programmes are introduced that are relevant to the operation of the program mechanisms" (p.6). Context considers not only locations, but interpersonal and social relationships, the biological, technological, and economic setting in which a program takes place. Outcome patterns (O) "comprise the intended and unintended consequences of programmes resulting from activation of different mechanisms in different contexts" (p. 7). A Context-mechanism-outcome pattern configuration (CMOC's) is a proposition stating what it is about a program, which works for whom in what circumstances (Pawson & Tilley, 2004). The purpose of a 'realist evaluation' then is to identify or test promising context-mechanism-outcome pattern configuration hypotheses. Qualitative and quantitative methods are employed to explain outcome footprints as likely results of mechanisms activated according to measures (interventions) in the context of the program and its participants.

Application of a ‘realist approach’ for an instructional coaching professional development evaluation is fitting due to the strong contextual focus. The coaching context by nature was very personalized and its strength was that the teachers received professional development in their individual setting, with their particular students and their particular lessons. It was less possible to control for variables within such a design than within a workshop model. Maxwell and Mittapalli (2010) cited Shadish, Cook, & Campbell (2002) stating, "that experiments do less well in clarifying the mechanisms through which and the conditions under which that causal relationship holds - what we call causal explanation" (p. 156). The following formula sums it up best: “mechanism + context = outcome" (Pawson & Tilley, 1997, p. xv) which placed causal explanation in a context dependency. (Maxwell & Mittapalli, 2010).

Several schools and districts implemented instructional coaching programs and designed or evaluated them (Friendship Public Charter School, 2010; Merola et al, 2011; APQC, 2011; Ithaca City School District Instructional Coaching Program, 2013). These practical program evaluations were larger scale evaluations by professional evaluation agencies using a variety of instruments such as surveys, interviews, and secondary source data. The studies aligned with instructional professional development programs that the district or school had invested in. However – these were practical evaluations without an explicit theory based approach and no intent to demonstrate an empirically valid effect. This professional development evaluation study used multiple data sources similar to the larger practical evaluations, grounded in program evaluation theory and focusing on causality in a complex context, and employed quantitative and qualitative methods. Semi structured interviews described and explained the instructional coaching professional

development from the views of the algebra teachers at the participating high school. Teachers were asked to describe their perceptions of professional development work and how it had affected their practice. The responses were coded for common themes to provide contextual explanations that helped shed light on the experience from the teachers' perspectives.

The study also analyzed written descriptions from participants in the form of learning logs and classroom observations using a researcher developed visitation protocol. The data was used to look for implementation of mathematical practices and teacher generated instructional strategies that were collected after the pilot year of the professional development implementation. Self-reporting formative and summative surveys were also conducted with the algebra teachers at the beginning and end of the 2014-2015 school year examining teacher learning, teacher practice, and student learning. A summative self-report student survey regarding student learning outcomes was conducted at the end of the 2014-2015 school year.

The quantitative analysis of formative and summative self-report surveys investigated the effects of the professional development on teacher practice and student learning from the teachers' perspective. The self-report summative student survey investigated the effect of the professional development on teacher practice and student learning from the students' perspective. Finally the qualitative responses provided further causal explanations about the mechanisms that lead to understanding the effects of teacher practice and student learning outcomes.

**Research Questions**

Overarching research questions:

1. Was it possible to create a 'realist evaluation' based framework to evaluate an instructional coaching model of professional development?
2. Realist Evaluation: What context-mechanism-outcome pattern configurations did the instructional coaching confirm to inform the further refinement of the professional development?
3. Exploratory Factor Analysis: Does the student survey measure teacher practice and student learning outcomes regarding the implementation of the Common Core State Standards for Mathematical Practice?

Guskey's (2000) five level framework for professional development evaluation was used to organize the evaluation questions per level:

1. How did the teachers describe their experience with the instructional coaching professional development (Guskey's levels 1 and 3)?
2. In what way did the instructional coaching professional development have an effect upon teachers' perceptions of their knowledge about teaching practices that implement the Common Core State Standards for Mathematical Practice (Guskey's level 2)
3. In what way did the classroom instructional coaching professional development have an effect upon teachers' perceptions of their use of teaching practices that implement Common Core State Standards for Mathematical Practice (Guskey's level 4)?

4. In what way did the instructional coaching professional development have an effect upon teachers' perceptions of their students' learning defined as the demonstration of skills that evidence the implementation of Common Core State Standards for Mathematical Practice in a classroom (Guskey's level 5)?
5. Teachers and students reported that instructional coaching professional development affected their use of teaching practices that implement the Common Core State Standards for Mathematical Practice (Guskey's level 4).
6. Teachers and students reported that instructional coaching professional development affected students' learning outcomes as measured by the demonstration of skills that evidence the implementation of Common Core State Standards for Mathematical Practice (Guskey's level 5)

Figure 7 below is a visual representation of the methods, conceptual framework and the worldview and data collection procedures that defined the methodological approach in this study with the overarching research question in the center.

Figure 7. Concept Map

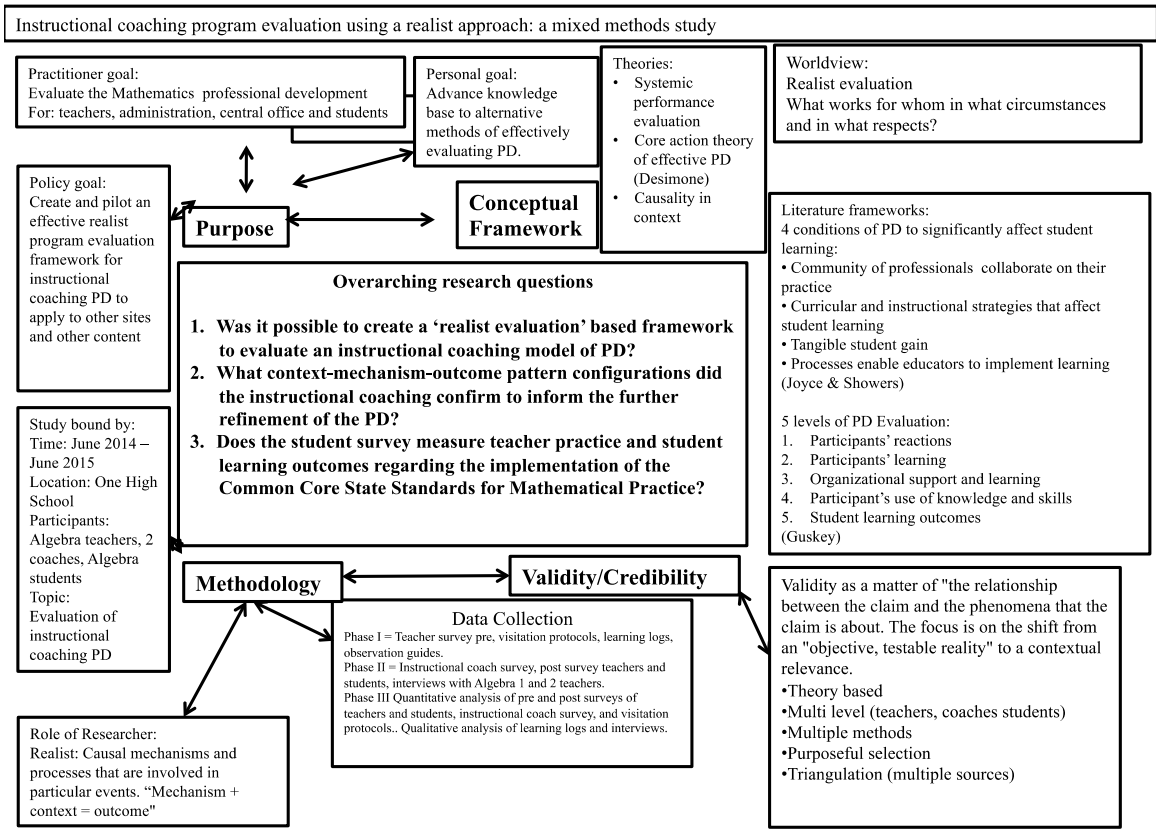



*Figure 7.* Concept Map. Adapted from "*Qualitative Research Design: An Interactive Approach" by J.* Maxwell, 2013. p.5. Copyright 2013 by Sage.

## Development of the Evaluation Framework

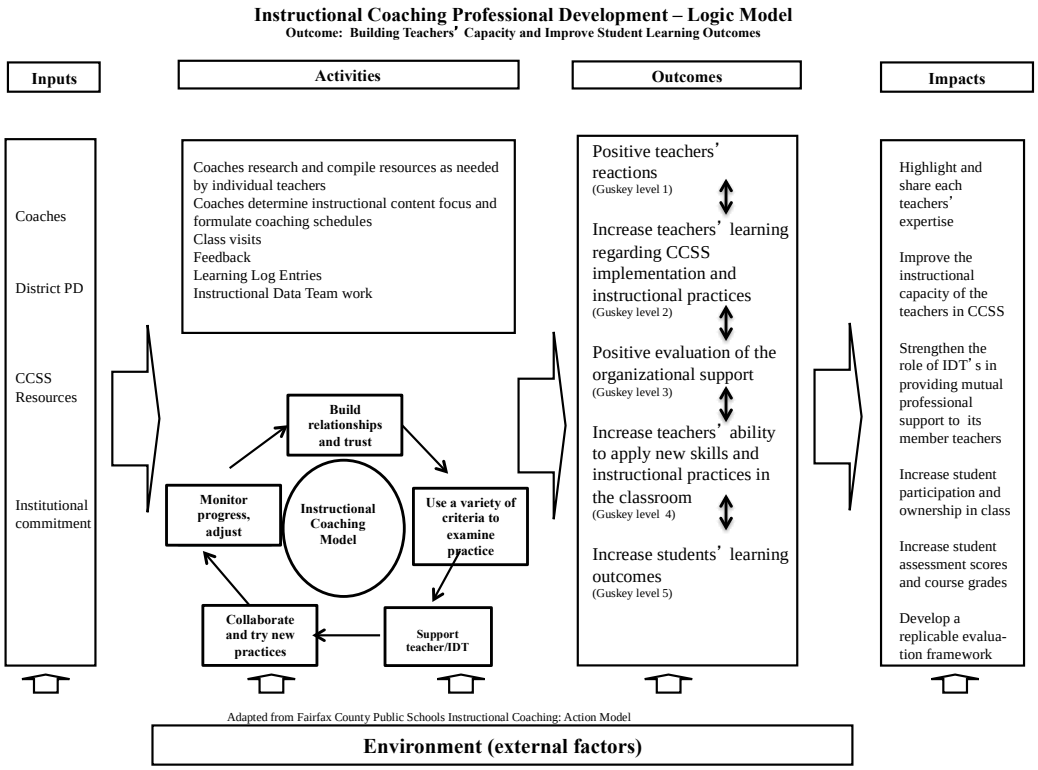
The purpose of this study was to develop a 'realist evaluation' based framework, to pilot the instruments that helped explain the mechanisms of professional development for mathematics. Theory-based evaluation required a program model to be developed that sequenced the causal path between inputs and intended outcomes (Bichelmeyer & Horvitz, 2006; Rogers et al., 2000). The theoretical assumptions served as a foundation for determining how the program was intended and expected to work, and was used to guide the evaluation processes (Rogers, Petrosino, Huebner, & Hacsi, 2000). Desimone (2009) recommended a conceptual framework that all empirical causal studies of

professional development should use. The model suggested an underlying core theory of action that included the following assumptions: Teachers experience effective professional development that increases their knowledge, skills, and attitudes to improve the content of their instruction which leads to increased student learning. Guskey (2000) developed five necessary levels that coincide with Desimone's (2009) conceptual framework that should be evaluated. They are participants' reactions, participants' learning, organization support and change, participants' use of new knowledge and skills, and student learning outcomes. This study applied the core conceptual framework that Desimone (2009) suggested and its theoretical assumption of interactive, non-recursive relationships between the critical features of professional development and teacher knowledge and beliefs, classroom practice, and student outcomes. The study followed the paradigm of a 'realist evaluation' as a way to measure effects and provide causal explanations through mechanisms that occurred in the context of the professional development and generated outcome patterns. The logic model for this study (see Figure 8) was developed to conceptualize the assumed causal paths between inputs, activities, intended outcomes, and impacts.

The activities of the coaches and teachers included Desimone's (2009) assumptions of content focus, active learning, coherence, duration, and collective participation. This required the provision of resources for teachers, the collaborative determination of instructional content focus, and setting up coaching schedules. The coaches visited classes, gave feedback, and continually tried to implement suggestions. The teachers filled out learning logs and collaborated in their instructional data teams.

The Logic Model (see Figure 8) visualized the assumption that these activities would lead to a positive experience for the teachers: That they would increase their knowledge and skills, and that that would lead to a change in instruction. That change in instruction would then positively affect student learning.

Figure 8. Logic Model




The Logic Model was developed within the larger scope of the ADDIE model (see Table 3) The use of the ADDIE model ensured the consideration and alignment of objectives of the professional development and the measures that would determine whether the goals were reached, at the planning stage (Haslam, 2010; Newcomer, Hatry, & Wholey, 2010).

Table 3

*ADDIE Model for Instructional Coaching Professional Development*

| Date | Process/Step | Key Activities |
|---|---|---|
| May - July 2014 | **Analyze**<br>Conduct needs analysis | • Collaborated with central office staff to determine district needs;<br>• Collaborated with site administration to determine perceived needs of staff<br>• Analyzed school performance results and areas of improvement |
| May - July 2014 | **Design**<br>Design the support model for the instructional coaching professional development | Researcher drafted **logic model** based on data analysis |
| May - July 2014 | **Develop**<br>Develop key work functions for the instructional coaching professional development | Collaborated with central office and site stakeholders to develop the key work functions.<br>• Coaches in collaboration with central office staff compiled resources as needed by individual teachers<br>• Coaches in collaboration with central office and researcher determined instructional content focus which included a schedule for coaches' visits<br>• Class visits<br>• Feedback<br>• Learning Log Entries<br>• IDT work |
| September 2014 - June 2015 | **Implement**<br>Implement the key work functions | Implemented the key work functions within the classroom embedded action model<br>• Build relationships<br>• Used a variety of criteria to examine practice<br>• Supported teacher/IDT<br>• Collaborated and discussed new practices<br>• Monitored progress, adjusted |
| June 2015 | **Evaluate**<br>Evaluate the impact of each key function | • Analyzed survey results, learning log entries, visitation protocols, semi-structured interviews, and student achievement data<br>• Presented findings to central office and determined conclusions and recommendations |

*Note:* Adapted from "The Addie Model: Designing, evaluating instructional coach effectiveness," by S. Danks, 2011, *ASQ Primary and Secondary Education Brief, 4(5*), p.2. Copyright 2011 by ASQ

The utilization of the ADDIE model (Danks, 2011) and Logic Model provided a systemic, contextual approach to evaluating instructional. The instruments were developed to measure and explain the mechanisms of the key work functions and their effects on outcomes. The utilization of the Logic Model ensured a systemic view of the instructional coaching and addressed the limitations of Guskey's (2000) level approach model by equating the levels with outcomes that the evaluation sought to explain. The systems approach (Laszlo & Krippner, 1998) shed light on the context in which the professional development took place and the Logic Model helped visualize the assumed causal connections that the instruments sought to measure and explain. The evaluation framework developed for this study was seeking to add to the knowledge of how theory based 'realist evaluation' using Desimone's (2009) theory of action could be used to provide replicable measures for alternative professional development designs such as instructional coaching as outlined by Guskey and Yoon (2009).

## Procedures

### *Role of the Researcher*

This study was based on a realist perspective. Maxwell and Mittapalli (2010) discussed realism as an integration of a realist ontology and a constructivist epistemology. There is a real world that exists independently, but our understanding of this world is the result of individual construction (Maxwell & Mittapalli, 2010). This connects well with mixed methods research that by definition includes qualitative and quantitative methods to answer a research question. This study is especially useful for proponents of the view that methods are naturally linked to paradigms as it highlights the

compatibility of various methods when some of the underlying paradigms are seen as conflicting. Mark, Henry, and Julnes (2000) state that the realist foundation of their evaluation theory offers "a satisfactory way to get beyond the paradigm wars" (p. 16). Quantitative and qualitative methods are seen "as assisted sensemaking techniques that have specific benefits and limitations" (p.16). Pawson and Tilley (1997) promote a realist perspective for evaluations and require a realist methodology in order to be an applied research "not performed for the benefit of science as such, but pursued in order to inform the thinking of policy makers, practitioners, program participants, and the public" (p. xiii). Maxwell and Mittapalli (2010) argued that critical realism is compatible with the characteristics of qualitative and quantitative research and enables communication between the two. They discussed four issues for which realism could make a contribution to mixed methods research: (a) A process approach to causality; (b) mind as part of reality; (c) validity and inference quality; and (d), diversity as a real phenomenon.

Within a realist approach, causality refers to causal mechanisms and processes that are involved in particular events. The mechanism is responsible for the relationship, not a variable but an account of the relationship, which opened the door for causation in qualitative research. The previously stated formula mechanism + context = outcome (Maxwell & Mittapalli, 2010; Pawson & Tilley, 1997) placed causal explanation in a context dependency. Realism then extended the concept of causation to an explanatory element of the context of the phenomena studied. Minds as part of reality addressed the fact that emotions, beliefs, and values, were also part of reality not just abstractions from behavior or constructions of the observer. Realism supported the notion that individual's social and physical contexts affected their beliefs and perspectives. Validity within

realism does not focus on procedures but was a matter of "the relationship between the claim and the phenomena that the claim is about" (p. 158). Again, the focus was on the shift from an objective, testable reality to a contextual relevance (Maxwell & Mittapalli, 2010). A realist perspective on validity could be very beneficial to mixed method researchers by focusing attention on particular threats to the conclusions drawn in a given study, which depend on the context and purpose of that study as well as on the methods used.

The researcher in this study holds an etic perspective in the sense that the relationship between researcher and participants was that of an evaluator of the empirical effects of the professional development as the participants perceived them (self-reporting surveys and learning logs) and as a direct observer of their behaviors (observation guides, visitation protocols). An emic perspective was also embraced in this study as the researcher shared the context in which participants received the professional development due to the administrative position the researcher held at the study site.

Maxwell and Mittapalli (2010) state that realist perspectives have a significant influence on mixed methods studies in program evaluation. Their examples illustrated the role of context-specific conditions that were only captured through deeper empirical probing and re-conceptualization. Critical realism in economics also provides a philosophical and methodological foundation for a broad set of alternative approaches. This researcher of this study applied the realist perspective to professional development evaluation as an alternative to either strictly experimental or quasi-experimental studies or qualitative studies due to the strong contextual setting of instructional coaching and the need for a process approach to causality (Maxwell & Mittapalli, 2010) to conduct an

effective evaluation. The realist perspective facilitates the practical purpose of evaluations to inform participants, practitioners and stakeholders (Mark, Henry & Julnes 2000; Pawson & Tilley, 1997).

### *Bounding the Case*

This study was conducted in a large school district in the Northeastern United States and was confined to a single High School. The research study was bound by subject area (Algebra 1 and Algebra 2) and included the teachers, their students and the coaches. Data collection occurred in the 2014-2015 school year. Eight Algebra 2 teachers and four Algebra 1 teachers (three were part of the pilot study, one wasn't) were included. The study concluded by the end of June 2015.

### *Data Collection*

The instruments were designed to collect data from each of Guskey's levels (2000). The quantitative instruments included surveys and visitation protocols (see Appendices B, D, F, and H respectively) that were designed to tally the frequency of observed practices. Semi-structured interview protocols and learning logs were used to gather the qualitative data (see Appendices C and E respectively). The instruments were selected or created based on general methods used in professional development evaluation (Guskey, 2000) and were specifically based on evaluations of instructional coaching programs (Friendship Public Charter School, 2010; Ithaca City School District, 2013). The development and design of the instruments followed Haslam's (2010) suggestion that the content of the instrument needed to be aligned with the content of the

professional development, and observation protocols should contain key elements that align with the professional development content. Haslam (2010) also suggests that classroom activities be based on well-defined rubrics and the surveys should include close-ended items aligned to the professional development content.

Instruments developed for this study were aligned to the Common Core State Standards for Mathematical Practice and the classroom observation guide developed by the GE Foundation. The observation guide was introduced to the district in previous years to support Common Core implementation. It listed instructional strategies and classroom activities that the observer would see teachers and students do if the mathematical practices were in place. Additionally the teacher and student surveys included strategies listed in *Instructional Implementation Sequence: Attaining the CCSS Mathematical Practices Engagement Strategies,* a document developed by LCM Math Leadership to support implementation of Common Core State Standards for Mathematical Practice.

Instruments (Appendix B-H):

#1: Instructional coach faculty survey (Appendix B: existing instrument, Alabama PEPE program):

The survey was adapted from an existing instrument and included the questions that referred to expectations the district had communicated to the coaches as far as their responsibilities were concerned. The adjustments were made in collaboration with the Director of School Improvement to ensure only items were contained in the survey that the district of this study was expecting from their coaches.

#2: Semi-structured interviews with teachers (Appendix C: researcher developed protocol):

The interview protocol mirrored a similar protocol used in the pilot study but organized by Guskey's levels of evaluation to provide a clear structure for the content analysis categories when analyzing the transcripts.

#3: Pre and post teacher survey (Appendix D: researcher developed instrument):

The formative and summative teacher surveys for Math classrooms were developed by the researcher in collaboration with the Director of School Improvement as well as reviewed by one mathematics teacher who participated in the pilot study to ensure the language and content reflected the district professional development vision and content and was clear to understand for teachers. The survey questions were constructed to directly align to classroom activities and actions of students and teachers according to the GE observation guides that the district had been using to promote CCSS implementation and the strategies listed in the implementation sequence of CCSS practices. The questions were organized in alignment with Guskey's Levels 2, teachers' knowledge and skills, Level 4, teachers' transfer of knowledge, and Level 5, student learning outcomes. The Likert scale responses selected to be appropriate to capture knowledge (Level 2; extensive, proficient, basic, no knowledge), transfer of knowledge and skills (Level 4; every lesson, regularly, sometimes, never), and student learning (Level 5; every lesson, regularly, sometimes, never). The connection between teacher knowledge and transfer to the classroom was captured through a tri-fold alignment of same item questions that is shown in Table 4.

Table 4

*Survey Question Examples Level 4*

| Teacher survey level 2 | Teacher survey level 4 | Student survey level 4 |
|---|---|---|
| **Describe your knowledge level regarding the following items:** | **How frequently do you engage in the following teaching strategies?** | **How frequently do you experience the following scenarios?** |
| The value and promotion of alternative approaches to solving problems; students are taught that mathematics is a sense making process for understanding.<br>☐ Extensive knowledge<br>☐ Proficient knowledge<br>☐ Basic knowledge<br>☐ No knowledge | I elicit, value, and celebrate alternative approaches to solving problems; students are taught that mathematics is a sense making process for understanding.<br>☐ Every lesson<br>☐ Regularly<br>☐ Sometimes<br>☐ Never | I am encouraged to use alternative approaches to solving problems; I am taught that mathematics is a sense making process for understanding.<br>☐ Every lesson<br>☐ Regularly<br>☐ Sometimes<br>☐ Never |
| The wording of the question was aiming to capture the degree of knowledge about this strategy not the degree of implementation in the classroom. | The wording of the aligned question in this section of the survey was aiming to capture the degree of implementation in the classroom from the teacher's perspective. | The wording of the aligned question in the student survey was aiming to capture the implementation of the strategy by the teacher from the student's perspective. |

Guskey (2000) stated that the key to gathering relevant information at Level 4 depended on the clear specification of indicators that revealed both the degree and quality of implementation. The surveys contained the specific practices that manifested CCSS implementation (quality) and the degree to which it was implemented from teachers' and students' perspectives. The construction of the survey items considered the assumed causal mechanisms occurring through the professional development process according to Desimone's (2009) core action theory to explain and measure the mechanisms and their effect (Pawson, 2013).

#4: Teacher learning logs (Appendix E: researcher developed protocol):

The teacher learning logs were developed to allow the teacher to write a journal entry regarding levels 2, 3, 4, and 5. The quarterly entries allowed the teacher to submit more recent recollections as the year progressed and written entries added another level of information to the semi-structured interviews at the end of the year.

#5: Visitation protocol Math (Appendix F: researcher developed protocol):

The visitation protocol was developed with a checklist in mind to provide the department head and administrator detail on what to look for when doing unannounced quick classroom visits. The checklist tallied the occurrences of practices during the visits over the duration of the study. The students' activities were based on the mathematical practices the district professional development has focused on aligned with CCSS implementation and that were reflected in the surveys. The teachers' activities in mathematics were generated from collaborating with the school improvement director and teachers participating in the pilot and in alignment with strategies contained in the surveys.

#6: CCSS Math classroom observation guide (Appendix G, Existing instrument, GE Foundation):

The observation guides have been developed by the GE foundation and made available for the district to support CCSS implementation. The have been given to administrators across the district but have not formally been utilized to visit classrooms and look for the particular practices and activities. The coaches, to guide their work with the teachers but not as a direct observation instrument, used them. They served as the foundation for the surveys.

#7: Summative student survey (Appendix H: researcher developed instrument, administered by the district and made available as secondary source data to the researcher):

The student survey followed the pattern in the teacher surveys and the questions were aligned to confirm or contest results from the teacher survey and to add the students' perspective (as shown in Table 5). The student surveys only contained questions regarding level 4 and 5.

Table 5
*Survey Question Examples Level 5*

| Teacher survey level 5 | Student survey level 5 |
|---|---|
| **I see the majority of my students do this during class:** | **I am doing this on a regular basis during class:** |
| Use varied representations and approaches when solving problems<br>☐ Every lesson<br>☐ Regularly<br>☐ Sometimes<br>☐ Never | I use varied representations and approaches when solving problems<br>☐ Every lesson<br>☐ Regularly<br>☐ Sometimes<br>☐ Never |
| The wording of the aligned question in this section of the survey was aiming to capture the degree of student learning from the teachers' perspective. | The wording of the aligned question in the student survey was aiming to capture the degree of student learning from the students' perspective. |

*Qualitative Data Collection:*

Semi-structured interviews with Algebra teachers were conducted in June of 2015. Additionally written descriptions from participants in the form of learning logs were collected throughout the school year, 2014-2015 on a quarterly basis.

*Quantitative Data Collection:*

The quantitative data collection for the professional development included an instructional coach survey and formative and summative self-report surveys in September of 2014 and May of 2015 with the Algebra teachers as well as a self-report summative survey with their students May 2015. It also contained visitation protocols and observations guides used in informal classroom visits to look for the implementation of mathematical practices and instructional strategies that were generated by the teachers after the first year of professional development in May 2014, in the form of frequency tallies.

*Sampling Criterion*

This mixed methods case study included all Algebra teachers, who received the professional development at the research site in 2014-2015. Survey results from all their students were included as well. The sampling scheme according to Collins (2010, p. 359) was "criterion" since the participating teachers had to receive instructional coaching professional development. Additionally the purpose of this study was to evaluate the professional development for the intended users. "Case studies become particularly useful when intended users (the district, school administration, and staff) need to understand a …program in great depth, and they can identify cases rich in needed information" (Patton, 2008; p. 458). The selection of the research site as opposed to all

three high schools in the district was a convenient sample. Convenience as a sampling scheme is applicable when groups or individuals are chosen, "who are conveniently available and willing to participate in the study" (Collins 2010: p. 359).

### *Data Analysis*

*Qualitative data analysis*

Interviews (instrument #2) with Algebra teachers were transcribed and coded for common themes and followed a content analysis (Hendricks 2013). The categories for coding were based on Guskey's (2000) evaluation framework and the five levels it contained, emerging themes within those categories followed Creswell's (2009) emerging themes. Learning log entries (instrument #4) were coded the same way and merged with the themes from the interviews. The responses were organized in a table by themes and summarized in a qualitative narrative connecting and discussing themes.

*Quantitative Data Analysis*

The analysis included descriptive statistics (mean and frequencies) because they were useful as a way of summarizing large number of scores in a single score (mean), and to display the results in a summarized form (frequency), (Norton 2009). Means from teacher and student surveys were compared. A paired t-test was conducted on pre and post survey scores to determine the statistical significance.

The quantitative data from the surveys (instrument #1, #3 and #7) were analyzed using means and frequencies as well as paired t-tests (#3). The results were presented in tables with the purpose to provide quantitative measures to assess the effect of the professional development on teacher practice and student learning. The visitation

protocol (instrument # 5) was analyzed using frequencies. An exploratory factor analysis was conducted for the student survey to determine what items of the survey would cluster together regarding a common factor.

*Data Collection and Analysis Plan for Classroom Embedded Instructional Coaching Professional development*

Instruments:
#1: Instructional coach faculty survey (existing instrument, Alabama PEPE program)

#2: Semi-structured interviews with teachers (researcher developed protocol)

#3: Pre and post teacher survey (researcher developed instrument)

#4: Teacher learning logs (researcher developed protocol)

#5: Visitation protocol (researcher developed protocol)

#6: CCSS Math classroom observation guide (existing instrument, GE Foundation, not directly collected but serves as a foundation for the other instruments)

#7: Summative student survey (researcher developed instrument)

Table 6 shows the data collection and analysis plan.

Table 6
*Data Collection and Analysis Plan*

| **Level of Evaluation** | **Data Source/Instrument** | **Data Analysis Plan** |
|---|---|---|
| Teacher's reactions (Guskey level 1) | Instrument #1:instructional coach faculty survey (existing instrument, Alabama PEPE program) | Statistical Analysis: Means |
| | Instrument #2: Semi-structured interviews with teachers (researcher developed protocol) | Creswell (2009) and Hendricks (2013): coding for emerging themes, content analysis |
| Teacher's learning (Guskey level 2) | Instrument #2: Semi-structured interviews with teachers (researcher developed protocol) | Creswell (2009) and Hendricks (2013): coding for emerging themes, content analysis |
| | Instrument #3: Formative and summative teacher survey (researcher developed instrument) | Statistical Analysis: Means, paired t-test |

| | | |
|---|---|---|
| | Instrument #4: Teacher learning logs (researcher developed protocol) | Text analysis to describe trends found in documents |
| Organizational support (Guskey level 3) | Instrument #2: Semi-structured interviews with teachers (researcher developed protocol) | Creswell (2009) and Hendricks (2013): coding for emerging themes, content analysis |
| | Instrument #4: Teacher learning logs (researcher developed protocol) | Text analysis to describe trends found in documents |
| Teacher's use of new knowledge and skills (Guskey level 4) | Instrument #2: Semi-structured interviews with teachers (researcher developed protocol) | Creswell (2009) and Hendricks (2013): coding for emerging themes, content analysis |
| | Instrument #3: Formative and summative teacher survey (researcher developed instrument) | Statistical Analysis: Means, paired t-test |
| | Instrument #4: Teacher learning logs (researcher developed protocol) | Text analysis to describe trends found in documents |
| | Instrument #5: Visitation protocol (researcher developed protocol) | Statistical Analysis: Frequencies |
| | Instrument #7: Summative student survey (researcher developed instrument) | Statistical Analysis: Means, exploratory factor analysis |
| Student learning outcomes (Guskey level 5) | Instrument #2: Semi-structured interviews with teachers (researcher developed protocol) | Creswell (2009) and Hendricks (2013): coding for emerging themes, content analysis |
| | Instrument #3: Formative and summative teacher survey (researcher developed instrument) | Statistical Analysis: Means, paired t-test |
| | Instrument #4: Teacher learning logs (researcher developed protocol) | Text analysis to describe trends found in documents |
| | Instrument #5: Visitation protocol (researcher developed protocol) | Statistical Analysis: Frequencies |
| | Instrument #7: Summative student survey (researcher developed instrument) | Statistical Analysis: Means, exploratory factor analysis |

*Note.* Adapted from "Application and Utility of the Guskey Professional Development Evaluation Model in a Community College Setting," by A. Ross, PhD diss., University of Tennessee, p.31. Copyright 2009 by A. Ross.

### *Data Merging*

The quantitative and qualitative data was analyzed for each of Guskey's (2000) levels. The results were merged per level to explain and measure the effects of the instructional coaching professional development on each of Guskey's levels. Data was merged per evaluation question not for each method separately (Greene, 2007).

## Credibility/Validity

The credibility and validity of this study was addressed through Patton's (2001) criteria set of evaluation standards and principles since the study was an evaluation. He referred to the Joint Committee on Standards for Educational Evaluation (2011), which encompassed four overarching criteria: Utility, feasibility, propriety, and accuracy. Patton (2001) concluded that those "who use evaluations apply both truth tests (Are the findings accurate and valid?), and utility tests (Are the findings relevant and useful?) (p. 550).

Utility addresses "the extent to which program stakeholders find evaluation processes and products valuable in meeting their needs" (Joint Committee, 2011; p.4). Evaluation process uses are benefits for participants because their participation in the evaluation process increased their own awareness or ownership (Joint Committee, 2011). The results of this study were presented to the district leadership and informed the design of the PD for the next school year. The participants were motivated to improve the PD design with their suggestions and ownership over their own learning was reflected in their goal setting meetings for the following school year.

Evaluations are feasible when they can be conducted effectively and efficiently (Joint Committee, 2011). The evaluation purpose and the instruments were introduced to

the district leadership as well as to the participating staff at the beginning of the school year. Participant's time and practicability were a key concern for the design of the instruments. Pilot study participants assisted and reviewed the instruments (survey questions, learning log questions and visitation protocol items) to ensure member input, meaningfulness and contextual viability (Joint Committee, 2011).

Propriety refers to "what is proper, fair, legal, right, acceptable, and just in evaluations" (Joint Committee, 2011; p. 106). Propriety was ensured through the Internal Review Board approval.

Accuracy addresses the truthfulness of evaluation representations. Accuracy is achieved through sound theory, methods, designs, and reasoning (Joint Committee, 2011). This study employed a theory based systemic approach to evaluate professional development (Desimone, 2009; Guskey 2000). "A well-conceived program theory provides a solid basis for interpreting evaluation findings and assessing causal claims" (Patton, 2008; p. 495). Evaluation tools such as the logic model and ADDIE model were used to develop the instruments and to ensure a systematic approach through alignment with the evaluation objectives as defined by the stakeholders. According to Patton (2008) emphasis is on methodological appropriateness – matching data collection and design to the nature of the evaluation and stakeholder priorities. The realist approach to analyze the results of this study ensured the appropriate consideration of the context while acknowledging observable regularities (Henry, Julnes & Mark, 1998).

In qualitative inquiry the researcher is the instrument (Patton, 2001). Patton concluded that therefore the experience, training and perspective of the researcher needed to be included in the study. The researcher of this study is an administrator at the study

site, and the direct supervisor of the participants. The Internal Review Board, due to the same interests of the researcher and the participants, did not deem it a conflict of interest or concern. The researcher and participants alike wanted to know if the district provided them useful professional development to implement the Common Core State Standards for Mathematical Practice. The researcher was a former Mathematics Department Head for two years and taught mathematics for four years. The researcher had conducted a pilot study the previous school year and interviewed the teachers participating in the pilot study. The results of those interviews guided the development of the semi-structured interview questions and fostered trust with the participating teachers of this study regarding the intent of the study to improve their professional development experience.

The credibility/validity of the instruments was addressed as follows:

#1: Instructional coach faculty survey (existing instrument, Alabama PEPE program, Appendix B): The survey was adapted from an existing instrument and reviewed by the school improvement director of the district to eliminate items that were not part of the coaches' services as negotiated between the district and the coaches.

#2: Semi-structured interviews with teachers (researcher developed protocol, Appendix C): Semi-structured interviews were conducted with four teachers participating in the pilot study the previous year as well as with the coach and the teacher on special assignment. Those interviews guided the development of the questions for this study's interview questions. The questions followed the program theory structure of Guskey (2000) and the responses were coded for Guskey's five levels and reported in the form of a narrative that included the frequency of same responses of teachers or tables including bulleted lists of items that were mentioned by the participant's.

#3: Pre and post teacher survey (researcher developed instrument, Appendix D): The survey included items that evidence the implementation of the eight mathematical practices of the Common Core Standards for Mathematical practice, instructional shifts, and student engagement as stated in the CCSS Math classroom observation guide (instrument #6). The items are grouped according to Guseky's (2000) levels teachers' learning (level 2), teacher's use of new knowledge and skills (level 4), and students' learning outcomes (level 5).

#4: Teacher learning logs (researcher developed protocol, Appendix E): The teacher learning logs included log entries regarding Guskey's levels 2,3,4 and 5. The entries were reported in the form of a narrative that included the frequency of same responses of teachers.

#5: Visitation protocol (researcher developed protocol, Appendix F): The visitation protocol contained the eight mathematical practices of the Common Core Standards for Mathematical Practice and a list of instructional strategies that was derived from instructional school improvement goals of the district and reviewed by the participants of the pilot study. The researcher and the current Mathematics Department Head conducted visits.

#6: CCSS Math classroom observation guide (existing instrument, GE Foundation, not directly collected but served as a foundation for the other instruments, Appendix G).

#7: Summative student survey (researcher developed instrument, Appendix H): The survey included items that evidence the implementation of the eight mathematical practices of the Common Core Standards for Mathematical practice, instructional shifts, and student engagement as stated in the CCSS Math classroom observation

guide (instrument #6). The items are grouped according to Guseky's (2000) levels teachers' use of new knowledge and skills (level 4), and students' learning outcomes (level 5).

**Validation of Student Survey (Exploratory Factor Analysis)**

An exploratory factor analysis was conducted for the student survey to validate the constructs of the student survey instrument. Williams, Brown, and Onsman (2010) stated that one of the applications of factor analysis was the provision of "construct validity evidence of self reporting scales" (p. 2). The survey intended to measure teachers' use of new knowledge and students' learning outcomes regarding practices that manifest the implementations of the Common Core State Standards of Mathematical Practice. The items were derived from the CCSS Math classroom observation guide developed by GE. The observation guide contained practices that evidence the implementation of instructional shifts (IS), mathematical practices (MP's) and student engagement (SE) when observing math classrooms. The survey contained 33 items. Questions 1-18 referred to teachers' use of instructional practices (as perceived by the students) and questions 19-33 referred to student learning outcomes related to these practices.

The factor analysis followed the five step protocol of exploratory factor analysis Williams, Brown, and Onsman (2010) listed (p.4): Is the data suitable (1); how will factors be extracted (2); what criteria will assist in determining factor extraction (3); selection of rotational method (4); interpretation and labeling (5).

*Is the data suitable for factor analysis?*

The sample size in this study was 359 submitted surveys wit 33 questions (variables), which constituted a good sample size as well as a good sample/variable ratio (see literature cited in Williams, Brown and Onsman, 2010).

*How will factors be extracted?*

In this factor analysis maximum likelihood was chosen as the mathematical method to get the loadings as Gorsuch (1989) recommended.

*What criteria will assist in determining factor extraction?*

In this factor analysis factors were extracted that had an Eigenvalue greater than 1 as well as based on the scree plot. The literature suggests to use multiple approaches in factor extraction (Williams, Brown, and Onsman, 2010).

*Selection of rotational method*

The rotational method that was selected was oblique rotation (direct oblimin) assuming that the factors are correlated, a reasonable assumption in social sciences in general and specifically regarding the theoretical constructs the survey was based on (teacher practice and student learning).

*Interpretation and labeling*

The factor analysis extracted three factors with an Eigenvalue greater than 1(see the total variance as well as the scree plot in Appendix I). The Pattern Matrix showed that factor 1 has 15 questions, which indicated a strong association between the variables and the factor, while factor 2 had 12 questions. Factor 3 did not show really strong associations. Five questions (6,13,14,15,16) showed no clear association. They were the questions pertaining to instructional shifts and student engagement (see Appendix I, Survey Questions). Question 2 was equally associated with factor 1 and factor 2 and with

low association. The survey questions with higher loadings in factor 1 were the questions pertaining to student learning outcomes and all the mathematical practices. The questions with higher loading in factor 2 were pertaining to student engagement and mathematical practice 1 in regards to teachers' practices. Factor 1 then was labeled student learning in the mathematical practices and factor 2 was labeled teacher practice that evidenced student engagement and the implementation of mathematical practice 1, "make sense of problems and persevere in solving them". The pattern matrix confirmed that the 27 questions with higher loadings for their respective factor were strongly associated with the constructs the survey intended to measure (see Table 7 and Appendix I, pattern matrix).

Table 7
*Pattern Matrix*

| Factor 1 Loadings | |
|---|---|
| Question 27 | .805 |
| Question 30 | .796 |
| Question 19 | .791 |
| Question 32 | .780 |
| Question 28 | .773 |
| Question 31 | .769 |
| Question 21 | .744 |
| Question 20 | .739 |
| Question 29 | .734 |
| Question 22 | .722 |
| Question 24 | .721 |
| Question 25 | .707 |
| Question 33 | .678 |
| Question 26 | .678 |
| Question 23 | .676 |

| Factor 2 Loadings | |
|---|---|
| Question 11 | .859 |
| Question 10 | .777 |
| Question 8 | .751 |
| Question 12 | .719 |
| Question 9 | .679 |
| Question 7 | .677 |
| Question 18 | .667 |
| Question 4 | .630 |
| Question 1 | .624 |
| Question 3 | .609 |
| Question 17 | .581 |
| Question 5 | .562 |

**Data Triangulation**

This study employed different sources and different methods to meet the validity strategy of triangulation. The instruments contained interviews with all participants, surveys of all teachers as well as surveys with students which increased the validity to the teacher responses by comparing similar teacher constructs with student constructs

responses. Both, qualitative and quantitative methods were used, namely interviews, as well as a surveys and learning logs and visitation protocols. The role of the researcher was clarified from the outset of the study. Creswell (2009) concluded that he recommended every researcher to engage in at least two of the before-mentioned strategies.

Validity within realism did not focus on the procedures alone but focused on the relationship between the claim and the phenomena that the claim was about (Maxwell & Mittapalli, 2010). This meant a shift from an objective, testable reality to a contextual relevance. The study's validity was based on the realist stance stated by Maxwell and & Mittapalli (2010) through the combination of quantitative survey data that sought to measure the effects of theory based mechanisms and the combination with qualitative data to contextualize the findings and focusing on plausible threats. The qualitative data clarified the why and how of the mechanisms through which and the conditions under which the effects were generated.

## Chapter 4: RESULTS

This study sought to evaluate the instructional coaching professional development that the district provided for the Algebra teachers by developing a realist framework based on Guskey's (2000) model for evaluation and Desimone's (2009) suggested core action theory of professional development. The process included the development, validation, and piloting of instruments. The instruments were constructed to align instructional practices that manifested the implementation of the Common Core State Standards for Mathematical Practice in classrooms with Guskey's (2000) levels of evaluation which include participants' reactions, participants' learning, organizational support, participants' use of new knowledge and skills, and student learning outcomes and in doing so providing explanations and measures of the mechanisms that lead to certain outcome patterns.

The evaluation questions (EQ) to evaluate the professional development were broken down into the following questions to directly address Guskey's (2000) levels of evaluation.

1. How did the teachers describe their experience with the instructional coaching professional development (Guskey's levels 1 and 3)?
2. In what way did the instructional coaching professional development have an effect upon teachers' perceptions of their knowledge about teaching practices that implement the Common Core State Standards for Mathematical Practice (Guskey's level 2)?
3. In what way did the classroom instructional coaching professional development have an effect upon teachers' perceptions of their use of teaching practices that

implement Common Core State Standards for Mathematical Practice (Guskey's level 4)?

4. In what way did the instructional coaching professional development have an effect upon teachers' perceptions of their students' learning defined as the demonstration of skills that evidence the implementation Common Core State Standards for Mathematical Practice in a classroom (Guskey's level 5)?
5. Teachers and students reported that instructional coaching professional development affected their use of teaching practices that implement the Common Core State Standards for Mathematical Practice (Guskey's level 4).
6. Teachers and students reported that instructional coaching professional development affected students' learning outcomes as measured by the demonstration of skills that evidence the implementation of Common Core State Standards for Mathematical Practice (Guskey's level 5).

The results of this study utlized six different instruments that addressed the five levels of Guskey's (2000) evaluation model to answer the program evaluation questions. The data collection and analysis plan in Table 8 provides an overview which instrument was used to collect data at what level of Guskey's evaluation model (2000).

Table 8

*Data Collection and Analysis Plan*

| Level of Evaluation | EQ | Data Source/Instrument | Data Analysis Plan |
|---|---|---|---|
| Teacher's reactions (Guskey level 1) | EQ 1 | Instrument #1:instructional coach faculty survey (existing instrument, Alabama PEPE program) | Statistical Analysis: Means |
| | | Instrument #2: Semi-structured interviews with teachers (researcher developed protocol) | Creswell (2009) and Hendricks (2013): coding for emerging themes, content analysis |
| Teacher's learning (Guskey level 2) | EQ 2 | Instrument #2: Semi-structured interviews with teachers (researcher developed protocol) | Creswell (2009) and Hendricks (2013): coding for emerging themes, content analysis |
| | | Instrument #3: Formative and summative teacher survey (researcher developed instrument) | Statistical Analysis: Means, paired t-test |
| | | Instrument #4: Teacher learning logs (researcher developed protocol) | Text analysis to describe trends found in documents |
| Organizational support (Guskey level 3) | EQ 1 | Instrument #2: Semi-structured interviews with teachers (researcher developed protocol) | Creswell (2009) and Hendricks (2013): coding for emerging themes, content analysis |
| | | Instrument #4: Teacher learning logs (researcher developed protocol) | Text analysis to describe trends found in documents |
| Teacher's use of new knowledge and skills (Guskey level 4) | EQ 3 EQ 5 | Instrument #2: Semi-structured interviews with teachers (researcher developed protocol) | Creswell (2009) and Hendricks (2013): coding for emerging themes, content analysis |
| | | Instrument #3: Formative and summative teacher survey (researcher developed instrument) | Statistical Analysis: Means, paired t-test |
| | | Instrument #4: Teacher learning logs (researcher developed protocol) | Text analysis to describe trends found in documents |
| | | Instrument #5: Visitation protocol (researcher developed protocol) | Statistical Analysis: Frequencies |
| | | Instrument #7: Summative student survey (researcher developed instrument) | Statistical Analysis: Means, exploratory factor analysis |
| Student learning outcomes (Guskey level 5) | EQ 4 EQ 6 | Instrument #2: Semi-structured interviews with teachers (researcher developed protocol) | Creswell (2009) and Hendricks (2013): coding for emerging themes, content analysis |
| | | Instrument #3: Formative and summative teacher survey (researcher developed instrument) | Statistical Analysis: Means, paired t-test |
| | | Instrument #4: Teacher learning logs (researcher developed protocol) | Text analysis to describe trends found in documents |

| | | |
|---|---|---|
| | Instrument #5: Visitation protocol (researcher developed protocol) | Statistical Analysis: Frequencies |
| | Instrument #7: Summative student survey (researcher developed instrument) | Statistical Analysis: Means, exploratory factor analysis |

*Note*. Adapted from "Application and Utility of the Guskey Professional Development Evaluation Model in a Community College Setting," by A. Ross, PhD diss., University of Tennessee, p.31. Copyright 2009 by A. Ross.

### *1. How did the teachers describe their experience with the instructional coaching professional development?*

Question one addressed the teachers' experience with the instructional coaching professional development. To answer this question the instruments were addressing two of Guskey's levels (2000), teachers' reactions to the professional development (level 1) and the organizational support teachers perceived to have (level 3). The instructional coach faculty survey, semi structured interviews with teachers and the teacher and coach learning logs (instruments #1, #2, and # 4) provided results to answer this question.

*Instrument #1: Instructional Coach Faculty Survey*

The coaches' survey was given to all twelve Algebra teachers. Three of the four Algebra 1 teachers have worked with their coach the previous school year as part of the pilot study. One Algebra 1 teacher worked with the coach the first year. All eight Algebra 2 teachers worked with their coach for the first year as part of the coaching professional development. The twelve teachers were familiar with their coaches from district professional development sessions in previous years.

Teachers rated the 22 statements (Table 1, Appendix J) on a Likert scale designed to measure the degree to which the coaches provided support to the teachers. One equated to rarely, two equated to sometimes, three equated to usually, and four equated to almost

always, and x equated “don’t know”. The items which received the lowest ratings pertained to supporting the teachers in; (a) establishing routines and procedures, (b) ensuring that standards and instruction were aligned, and assistance in instructional planning was offered when needed. Items which received the highest ratings included scheduled meetings and the provision of feedback (x was excluded from the calculation of the mean).

There was a difference between the mean score of the Algebra 1 teachers (3.33) and the Algebra 2 teachers (2.68). Additionally the Algebra 1 teachers who were working with the coach the second year had an average rating of 3.68. Overall the mean score for the coaches was 2.88. Algebra 1 teachers rated their coaches performance overall higher than Algebra 2 teachers due to the relationship they had formed the previous year.

The mean scores from each teacher (Table 2, Appendix J) showed one teacher with a significantly lower score (1.32) then the next lowest ones (2.00 and 2.23). All three lowest mean scores were from Algebra 2 teachers. The highest mean scores (4.0, 3.64 and 3.57) were from two Algebra 1 teachers who had worked with the coach the previous year and one Algebra 2 teacher who had worked with their coach for the first year. The low score from one teacher suggests that the teacher did not experience the instructional coaching PD as it was intended.

*Instrument #2: Semi-structured interviews with teachers*

All participating Algebra teachers were interviewed at the end of the school year 2014-2015 to gain insight and provide explanations beyond the survey data since every participant formed their own relationship and worked individually with the instructional coach. The first set of questions addressed the participants experience with and reaction

to the instructional coaching professional development. The participants were asked to recall their prior professional development experiences in the district to help clarify the context of instructional coaching professional development.

*How has district Math* professional development *been designed prior to this instructional coaching* professional development*?*

All teachers reported that the previous professional development was usually a lecture session aligned to Algebra 1 and Algebra 2 teachers. While teachers pointed out that they had the opportunity to meet all the other math teachers in the district, they perceived that the professional development was irrelevant to their teaching in general and not geared towards their needs.

*Did you ever take something away from it that you were then implementing on a regular basis?*

Four teachers remember the Laying The Foundation (LTF) professional development and that they included the activities or the approach to learning mathematics (requiring students to learn the "why") in their instruction. Two teachers indicated that they used one or two activities and three teachers indicated they implemented nothing they learned about in previous professional development training.

*What were the strengths and weaknesses of that* professional development*?*

Five teachers reported it was beneficial to get together with peers to exchange ideas and thoughts. However, nine teachers perceived the professional development as irrelevant to the content or to the teaching practices for their classroom.

*How was the instructional coaching professional development different?*

All teachers reported that the professional development was individualized and personal. Two Algebra 2 teachers felt it was not beneficial for them and both suggested conferring with the coach before teaching the lesson. Ten teachers perceived the professional development as very beneficial, teacher specific, and relevant to their teaching.

*What about the instructional coaching* professional development *would you continue to do and why? What would you change and why?*

All teachers reported that they would like to continue the instructional coaching professional development. The Algebra 1 teachers, who have done the coaching for two years, would like to broaden it, add new challenges and involve peer coaching or a different person with a new perspective. Three of the Algebra 2 teachers would consider a different coach and three Algebra 2 teachers expressed an interest in having pre- instead of post- observation conferences. One teacher would like assistance with assessment development, and one would like more demonstration lessons. Two teachers stated they would like the coach to come less often and two teachers felt the visits should be more at the beginning of the year and not in the last quarter.

In regards to the organizational support that teachers perceived they received, they were asked the following questions:

*Have you felt that the organization has supported the instructional coaching professional development, the department head, administration, scheduling, or were they barriers to do this?*

All teachers reported that they felt supported. One teacher would have liked specific support in the form of technology resources. One teacher stated that the

consideration of class sizes would have helped to support the professional development and to schedule smaller class sizes for Algebra 1 and Algebra 2 in the future. One teacher reported that their requests to have a demonstration lesson were met which made them feel supported. Another teacher noted that he felt safe trying new things out because of the support of the department head and administration of the instructional coaching professional development.

*How has your collaborative team (IDT team or cluster) positively or negatively influenced the professional development?*

All teachers reported that the IDT's were not a hindrance but did not support the professional development. In four cases, strategies were discussed or shared. One teacher reported that he tried to share strategies with IDT members who were not part of the professional development.

*How has scheduling influenced the success or challenge with the professional development?*

Six teachers answered this question specifically and felt the scheduling was very flexible and accommodating (the other six teachers covered the topic of scheduling in the general question of organizational support and the interviewer did not ask this specific question). Two teachers mentioned that the debriefing during another class period however was challenging. One teacher mentioned that when coaches come in the first period or before a break the students were more difficult.

*Instrument #4: Teacher and coach learning logs*

The purpose of the learning logs was to collect qualitative data as to what teachers learned, in what way they felt supported or not supported by the organization, what they

implemented, and if they saw their students change as a result of the instructional coaching professional development. To get more detailed reports over time than from the semi-structured interview at the end of the school year, teachers submitted the learning logs after the first, second, and third quarter.

Five teachers felt that the IDT did not support their work because the time was too short to do meaningful work or the membership was too small (only two members in the Algebra 1 IDT's). Four teachers felt their IDT supported them. One teacher stated that a lot of information that enhanced student learning was shared at their IDT meetings and another felt that their IDT was helpful to discuss curriculum issues, as well as teaching strategies. One teacher reported that the weekly conversations at their IDT meetings helped improve their instruction and also student performance especially on their IDT created post tests.

Four teachers reported explicitly that the work of the coaches supported them due to their suggestions and review of their classes. They were given specific ways in which to modify particular lessons to give students more clarity. One teacher stated that the coach's perspective gave him/her more ideas how to improve the lesson. One teacher stated that the visit to another teacher's class when the coach modeled a lesson was most useful.

Half the teachers reported that they felt supported by the administration and the other half stated that they felt supported by their peers. Two teachers stated that their peers shared ideas with them.

One teacher felt supported in the use of IXL while she was not supported in the use of technology such as I-pads. One teacher stated that building maintenance did not

support him/her due to their unwillingness to make changes that would improve his ability to use the promethean board. One teacher stated that she was not supported by the district professional development because the topics were not relevant for her and she has no access to the manipulatives and software that were introduced. One teacher stated that the work on the curriculum committee with a colleague from another high school had shown him some insights on teaching why and reasoning in some topics.

The teachers interpreted the question on the learning log regarding organizational support very differently. Some listed any item or issue in regards to their instruction while others responded in regards to the instructional coaching and the support or lack of support they received.

In summary teachers perceived the instructional coaching professional development as a positive experience and felt the coaches were providing the support they were asked to provide. The Algebra 1 teachers who had previously worked with their coach rated their coach higher than the Algebra 2 teachers regarding the different supports the coaches were providing. All teachers reported that the professional development was individualized and personal, and they would continue the instructional coaching with suggestions to the frequency of visits, timing of the conference (pre instead of post) and the involvement of peers to observe one another. All of the teachers perceived the professional development as very beneficial, teacher specific and relevant to their teaching. Two teachers felt it wasn't beneficial for them. Three teachers would consider a different coach.

***2. In what way did the instructional coaching professional development have an effect upon teachers' perceptions of their knowledge about teaching practices that implement the Common Core State Standards for Mathematical Practice?***

Program evaluation question two addressed the teachers' perception of their knowledge about teaching practices that implement Common Core State Standards for Mathematical Practice before and after the instructional coaching professional development. Instruments #2, #3, and #4 provided results to answer this question.

*Instrument #3: Pre and post teacher survey*

All twelve teachers were given the same pre and post survey at the beginning and at the end of the school year 2014-2015. The survey was created to align with Guskey's levels in order to measure teacher's perceptions of their knowledge about teaching strategies compared to their perception of implementing these strategies and then students' learning outcomes as a result of those strategies. The first part of the survey pertaining to knowledge contained 19 items. Teachers were asked to rate their knowledge about a certain strategy on a Likert scale with 1 equaling no knowledge, 2 equaling basic knowledge, 3 equaling proficient knowledge, and 4 equaling extensive knowledge. The 19 items were taken from the CCSS Math classroom observation guide (instrument #6, Appendix G). This observation guide was developed by the GE foundation in collaboration with Stamford public schools and described what teachers and students would be doing if the implementation of CCSS mathematical practices was in preparation, getting started, moving along, or in place. The items contained in the survey are listed under the "in place" rubric of the CCSS Math classroom observation guide (Appendix G).

The mean scores ($\bar{x}$) of the teacher self-ratings in the pre and post survey (table 1, Appendix J) show, that teachers perceived their knowledge about teaching practices that evidence CCSS mathematical practices are implemented in a classroom on average at 2.94 at the beginning of the school year (2 equaling "basic knowledge" and 3 equaling "proficient knowledge") and at the end of the school year teachers' perception of their knowledge increased to an average of 3.25, an increase of 10.5 %. The perception of their knowledge increased for all items over the school year. The items with the largest increase included students' choice and the availability of multiple activities as well as the provision of multiple representations to support visualization of skills and concepts.

*Instrument #4: Teacher and coach learning logs*

The learning logs in research question two focusing on program evaluation used qualitative data to ask what teachers learned as a perceived result of the instructional coaching professional development. To get more detailed reports over time than from the semi-structured interview at the end of the school year, teachers submitted the learning logs after the first, second, and third quarter.

Nine teachers reported on learning more about group work. They learned about using assigned roles, and grouping students purposefully based on mixed skill levels, and how it promoted student centered learning. They knew more about engaging all students within groups in order to encourage all students to contribute their ideas, feel a responsibility to understand all concepts and want to not let the group down.

Two teachers learned about higher order questioning and to focus on Depth of Knowledge (DOK). Two teachers reported that they learned more about making students think conceptually. Additionally, three teachers listed learning about promoting students'

independence, using strategies such as “ask three before me” and waiting for students to struggle more before giving them answers. One teacher mentioned the use of online resources to improve prerequisite skills, current content, and challenging those that finish early. One teacher listed creating assessments that focus on real-life application problems and grading with a four point rubric. Three teachers listed in the latter quarters of the year that they gained knowledge regarding the use of manipulatives to increase conceptual understanding of students as well as designing discovery based lessons. Two teachers mentioned the model lesson the coach taught and noted how it taught them a new way of introducing a topic as well as discovery based activities.

One teacher reported that s/he did not learn anything about new strategies.

The Algebra 1 coach reported that teachers learned more about paying attention to structure (Mathematical Practice #7 and Mathematical Practice #8), and about embedding the big ideas of the unit in the lesson. The Algebra 2 coach reported that teacher opened up in regards to student centered lessons instead of focusing on the instructor and to develop lessons that lead to students making mathematical connections rather than conveying formulas and procedures.

*Instrument #2: Semi-structured interviews with teachers*

The second set of questions addressed the participants’ perception of knowledge what they gained through the instructional coaching professional development. They were then asked what they would like to learn more about the following year. Tables 9 and 10 provide a summary of the items teachers mentioned they learned (Table 9) and what they would like to learn more about (Table 10).

*What have you learned through the instructional coaching professional development about teaching practices that implement CCSS?*

Table 9
*Teacher's Knowledge*

| **Algebra 1** |
|---|
| • Group work, improved scores across the board.<br>• Group work, students have assigned roles; student to student discourse.<br>• Give students time to answer and make them work instead of the teacher doing all the work.<br>• Group work with individual responsibilities. |
| **Algebra 2** |
| • Challenge the thinking of students, discovery based.<br>• Discovery based introduction of a particular topic.<br>• Managing transitions and setting the stage for what was to come; shorter types of investigations and how to use group work for that.<br>• Observing a demo lesson and see a simple introduction of a new topic that was student centered.<br>• Be less procedural and challenge the students to think about the why.<br>• Non-procedural focus, and attending to precision by using the correct mathematical characters.<br>• Allow for student discussion.<br>• Grouping, non-procedural focus when teaching for example factoring (see quote). |

One Algebra 2 teacher summarized his learning the following way:

> "The coaches big thing was to doing things in such a way that students realize the number sense of what they are doing and if you understand the conceptual nature behind what it is that you are doing the procedure will just come through that".

*What would you like to learn more about in a professional development next year?*

Table 10
*Changes for Next Year's Professional Development*

| |
|---|
| • Deciding what criteria to consider when grouping.<br>• Higher order questioning strategies.<br>• Refining group work and how to move students from one group to another periodically.<br>• Real life applications and interdisciplinary content between math and science.<br>• Visual strategies.<br>• More about the use of group work.<br>• One-day projects.<br>• The use of technology.<br>• More student directed teaching strategies. |

In summary teachers reported through the survey, learning logs, and semi structured interviews that they perceived their knowledge about teaching practices that implement the mathematical practices of CCSS had increased through the instructional coaching professional development. The survey items with the largest increase, were students' choice and availability of multiple activities as well as the provision of multiple representations supported the goal of the professional development to increase student centered learning in math classrooms. The development of number sense aligned with the more conceptual instead of procedural focus to increase understanding.

***3. In what way did the classroom instructional coaching professional development have an effect upon teachers' perceptions of their use of teaching practices that implement the Common Core State Standards for Mathematical Practice?***

Program evaluation question three addressed the teachers' perception of their use of teaching practices that implement Common Core State Standards for Mathematical Practice before and after the instructional coaching professional development. Instruments #2, #3, #4, and #5 provided results to answer this question.

*Instrument #3: Pre and post teacher survey*

The second part of the survey pertaining to the use of teaching practices contained 24 items. Teachers were asked to rate the frequency of engaging in a certain teaching practices on a Likert scale with 1 equaling never, 2 equaling sometimes, 3 equaling regularly, and 4 equaling every lesson. The 24 items were taken from the Instructional Implementation Sequence: Attaining the CCSS Mathematical Practices Engagement Strategies (LCM 2011), and the CCSS Math classroom observation guide (instrument

#6). The mean scores of the teacher results self-ratings in the pre and post survey (Appendix K) show that teachers perceived the frequency of their use of teaching practices that evidence that Common Core State Standards for Mathematical Practice are implemented in a classroom on average at 2.85 at the beginning of the school year (2 equaling "sometimes" and 3 equaling "regularly"). At the end of the school year teachers' perception of their use of these items was rated on average 3.06, an increase of 7.4 %. The perception of their use increased for all but three items over the school year. The items, which showed a decrease were, the effectiveness of class routines, students' choice in how to demonstrate their learning, and the use of formative assessments.

A decrease could indicate a validity concern since knowledge about a teaching practice can't really decrease. Teachers might not remember how they rated themselves at the beginning of the year and the time span between pre and post was 10 months.

*Instrument # 5: Visitation protocol*

The visitation protocol contained a checklist of skills that students would demonstrate in the classroom (the eight mathematical practices of the CCSS) and six teaching practices that teachers would use. These teaching practices were identified in collaboration with the school improvement director, and the teachers participating in the pilot study.

During the course of the 2014-2015 school year the assistant principal (researcher) and the department head recorded a total of 33 visits. During the short, informal, unannounced visits, they checked off items that were observed during the time of the visit. The frequency of the instructional strategies used by the teachers (Appendix

L) simply serves as a description of frequency of occurrence this school year during unannounced visits since there was no baseline to compare it to.

All selected teaching practices were observed in at least 27% of the visited classes. In 52% of the visited classes the teacher modeled, in 42% of the visited classes higher level prompting questions were asked, and in 33% of the visited classes student-to-student discourse was promoted.

*Instrument #4: Teacher and coach learning logs*

The third question on the quarterly learning log asked the teachers and the coaches what the teachers were implementing. Nine teachers stated that they implemented group work. One teacher grouped the students according to their skill levels. The teacher noticed that prior to having students in groups some students were much less engaged than when they sat in groups. One teacher reported that s/he engaged all students within the group in order to encourage all students to contribute their ideas, feel a responsibility to understand all concepts and want to not let the group down. One teacher used peer tutoring within group work. And another teacher taught the students how to discover material in cooperative groups without giving them a process to follow first. Four teachers used explorations or student centered activities to promote collaboration and give students the opportunity to demonstrate their learning in a project. Two teachers reported that they made students think conceptually. Three teachers promoted students' independence and waited for students to struggle more before giving them answers. One teacher promoted more "peer talk" with regards to creating solution tactics for problem situations. Two teachers implemented strategies they observed from the coach's model lesson and one of them requested a copy of the Algebra 2 coach's

model lesson on rational functions. He modified it and re-taught his Algebra 2 class based on the tips he had learned from the coach. One teacher added more emphasis to vocabulary and verbal explanations on simple questions, especially on assessments.

The Algebra 1 coach reported that in 38% of the classes observed, the primary activity consisted of students conducting an investigation, formulating conjectures, and answering probing questions. He observed one teacher implement for the first time assigning specific roles with clear expectations to members of cooperative groups. He saw three different teachers ask students to examine the structures of equations or inequalities (comparing and contrasting) and connect them to solution patterns. He has often seen teachers use "Do Nows" as a means to activate prior knowledge to help the students meet lesson targets. The Algebra 2 coach stated that teachers were using group work more successfully encouraging students to answer each other's questions instead of answering them themselves. Teachers were creating more interactive activities for students to work through and were connecting prior concepts to new concepts more regularly.

*Instrument #2: Semi-structured interviews with teachers*

The third set of questions addressed the participants' perception regarding the implementation of strategies on a regular basis as a result of the instructional coaching professional development. They were then asked if they thought that the instructional coaching professional development strengthened the teaching practices that implemented Common Core State Standards for Mathematical Practice.

*What instructional strategies have you implemented on a regular basis as a result of the instructional coaching professional development?*

Table 11 lists the strategies teachers reported that they have implemented on a regular basis due to their work with the coach.

Table 11
*Teachers' Use of Knowledge and Skills - Implementation*

- Grouping.
- Group work, student discourse and perseverance.
- Group work, perseverance, student accountability.
- Group work, majority of the lesson focused on problem solving, perseverance.
- Box and Whiskers activities that were discovery based and student centered.
- Discovery based activities to introduce a new topic and activities for box and whisper plots shown in a demo lesson.
- Transitioning, making expectations clear for different parts of the lesson, questioning strategies, student ownership.
- Group work.
- Partnered seating, student-to-student discourse as a consequence of it.
- Creating a friendly classroom environment and getting distracted students to focus. Particular design of a lesson that was demonstrated and teacher re-taught.
- Purposeful grouping and higher order thinking questions that pertained to conceptual understanding instead of a procedural focus.

Six teachers responded that they implemented group work, which fostered student engagement, accountability, responsibility and perseverance as well as higher order thinking and conceptual understanding. Two teachers mentioned more discovery based activities.

*What do you think about the instructional coaching professional development as a way to strengthen implementing Common Core State Standards for Mathematical Practice?*

Ten of twelve teachers perceived that the instructional coaching clearly strengthened implementation of strategies due to the one on one approach with immediate feedback and the relevance to the teachers' classroom. Two teachers perceived that the coaching could strengthen implementation but they had not perceived it as successful as it could have been for them. Both teachers did not feel they had a very positive experience with the coach at the beginning of the school year and felt the feedback was

too critical and too detailed. Teachers were able to make suggestions to the coach, as the school year progressed, to which s/he adhered. One request included the coach giving a demonstration lesson, which both teachers felt was very beneficial. Another teacher mentioned the provision of useful student centered activities that the teacher successfully implemented.

One Algebra 2 teacher described his experience with the coach the following way:

> "Because, if I go to a one-time workshop and I hear a brilliant speaker talk about something and he gives me great ideas then I go back to my classroom or back to my house and I'm thinking about it, I will try it once but then I will never see this person again or do not interact one on one, whereas, when I get something from our coach I can try something and then we can talk about it and then we can have a lot of follow up. So the follow up is a good thing".

In summary teachers reported in the survey learning logs, and semi structured interviews that they perceived their use of teaching practices that implement Common Core State Standards for Mathematical Practice has increased through the instructional coaching professional development. The teachers reported using twenty-two of twenty-five teaching practices in the survey more often at the end of the professional development than before the training. The three practices with the largest increase were encouraging students to explain their reasoning, maintaining high cognitive demand throughout lessons, requiring students to deeply engage with making sense of the mathematics, and justifying their thinking and differentiation so all students work at an appropriate level. In the learning logs and in the semi structured interviews, teachers

reported that they implemented group work and provided more student centered activities to engage students and promote student to student discourse. The results from the visitation protocol confirmed the use of these practices on a regular basis in the classrooms. The one on one format with immediate feedback was mentioned by ten of twelve teachers as the reason why instructional coaching professional development strengthened implementation. Nearly every item on the learning log that teachers stated to have learned they also reported they implemented.

***4. In what way did the instructional coaching professional development have an effect on teachers' perceptions of their students' learning defined as the demonstration of skills that evidence the implementation of the Common Core State Standards for Mathematical Practice in a classroom?***

Program evaluation question four addressed the teachers' perception of their students' learning defined as the demonstration of skills that evidence the implementation of Common Core State Standards for Mathematical Practice as a result of the instructional coaching professional development. Instruments #2, #3, #4, and #5 provided results to answer this question.

*Instrument #3: Pre and post teacher survey*

The third part of the survey pertaining to student learning outcomes contained 15 items. Teachers were asked to rate the frequency with which they saw students demonstrate 15 different skills on a Likert scale with 1 equaling never, 2 equaling sometimes, 3 equaling regularly, and 4 equaling every lesson. The 15 items were taken from the GE CCSS Math classroom observation guide (instrument #6). The survey items

were descriptions of what students would do from the rubric "in place" regarding the mathematical practices. The mean scores of teacher ratings in the pre and post survey regarding (table 3, Appendix K) the observed student learning as well as the difference between the means show the teachers' perception of student demonstrating items pertaining to Guskey's (2000) level 5, were on average at 2.59 at the beginning of the school year (2 equaling "sometimes" and 3 equaling "regularly"). At the end of the school year teachers' perception of students' frequency with which they demonstrated the selected skills was rated on average 2.81, an increase of 8.6 %. The perception of their students' demonstration of skills increased for all but one item over the school year. The items, which showed the lowest difference between pre and post measure of student learning outcomes were the application of knowledge and making assumptions, solving everyday life problems, and the analysis of mathematical relationships. The three items with the highest increase were the use of varied representations and approaches when solving problems, the maintaining of oversight of the mathematical process while attending to detail, and the use of mathematical terminology and vocabulary with precision.

*Instrument # 5: Visitation protocol*

During the course of the 2014-2015 school year the assistant principal (researcher) and the department head recorded a total of 33 visits. The frequency of the skills students demonstrated during the classroom visits (Appendix L) simply serves as informational evidence of implementation since there was no baseline for frequencies from a previous year.

All mathematical practices were observed in at least 15% of the visited classes. In 48% of the visited classes the students attended to precision, in 45% of the visited classes the students modeled with mathematics, and in 36% of the visited classes students were making sense of problems and persevered in solving them.

*Instrument #4: Teacher and coach learning logs*

The fourth question on the quarterly learning log asked the teachers and the coaches in what areas they perceived their students had improved.

Nine teachers reported that their students increased in their perseverance to solve problems. Three teachers reported that their students' improved their ability to learn cooperatively. Six teachers stated that the student-to-student discourse increased. Five teachers reported that students improved in solving real world problems. Two teachers reported that their students were more engaged and more independent. Three teachers reported students' improvement on graphing functions. One teacher noted that students improved their ability to explain their thinking, and one teacher reported that they wanted to know more about the why. One teacher reported that there was no observable improvement of students due to the instructional coaching professional development. The Algebra 2 coach reported that students seemed more comfortable in their groups, and were better at working together productively. The Algebra 1 coach saw students making progress in their perseverance in solving problems and exploring challenging concepts. He also saw them becoming more comfortable in formulating their own conjectures rather than expecting teachers to provide them with a set of rules to follow. He did however not feel that he had enough information to determine whether students have made significant improvement with fluency in the skills of algebra.

*Instrument #2: Semi-structured interviews with teachers*

The last question teachers were asked in the semi-structured interviews was pertaining to their students' learning. Table 12 lists the responses to the following question.

*How do you think has the instructional coaching professional development affected student learning?*

Table 12

*Student learning outcomes*

| |
|---|
| • Working in groups with assigned roles and talking math, explaining their solutions and compete with each other.<br>• Group work fostered student-to-student discourse and conceptual understanding evidenced by what the students were saying. More student engagement.<br>• More student-to-student discourse, better conceptual understanding which was also reflected in better scores. More student engagement.<br>• More student perseverance, more conceptual understanding and discourse about math. More student engagement.<br>• More student engagement, more conceptual understanding.<br>• A couple of things were independently successful but overall no.<br>• Students were more self-reliant. (Evidenced by what they were saying to each other – "I hate when she does that. She just answers your question with a question!")<br>• Students were more comfortable working in groups.<br>• Students worked better together in groups, their conceptual, non procedural learning has increased evidenced by observation but difficult to measure.<br>• Students were more independent and more engaged. But not all of them.<br>• Students worked better in groups which lead to more student engagement. Better conceptual understanding, which lead to better retention of knowledge as observed when reviewing material for finals. |

Ten out of twelve teachers perceived increased student learning as a result of the instructional coaching professional development. Five teachers reported that students worked better in groups and increased and improved their discourse. Six teachers specifically mentioned that students were more engaged and six teachers stated that students had a better conceptual understanding as a result of instructional changes. Three teachers mentioned more perseverance or independence.

One Algebra 1 teacher shared one student's learning experience:

"I think that it also helps …..I had a kid this year; I nominated her for the last dream award and she just did awesome. She said she's never had above a C in math in her life, she hated it, she said she sucked at it. She was in math center, which I think helped her a lot to get confidence in herself and then in my room she was like a little leader. She would walk around and help kids and I mean you know there are the studies out there that if you explain something you learn it better than if you just listen. I think she ended up with an 85 on her final and she was mad. She just came in my room to find out her grade and she was mad that it wasn't higher than that but I mean an 85 is awesome for someone that struggled. I think it helped her a lot to be the one that went around and explained, to defend her answers and say "no I did it this way" and she was always willing to help people out. She could've been difficult if it was traditional roles. She had a lot of energy, always drinking coffee, so I think she could've been a challenge but instead she ended up being the class leader."

When discussing her overall improved results at the mid terms and finals she explained:

"I don't know. I think them having to talk math. Because our final is a little different than it has been in the past. There are more questions that say "explain this…" - it's different. There was one where the question was solved incorrectly and it asked where did this girl make the mistake? And they had to explain. And when I got to those questions I was shocked at how well my kids explained what they were doing so I think it's just them having to explain every day… that made them able to do that."

Another Algebra 1 teacher reported:

> "I think it has definitely affected their perseverance, the willingness to participate, question, you know, themselves, me, each other within the classroom. You know for the most part there was participation in the classroom.…..Comparing last year's Algebra classes to this year's Algebra classes and you know how discipline issues that never resolved themselves last year for the most part were taken care of before Christmas this year. My not having to be right on top of someone every single minute in order to get them to do something, that improved."

In summary teachers reported through the survey, learning logs and semi-structured interviews that in their perception the instructional coaching professional development had a positive effect on their students' learning defined as the demonstration of skills that evidence the implementation of the mathematical practices in the classroom. Teachers rated the frequency by which certain skills were demonstrated on average at 2.59 at the beginning of the school year (2 equaling "sometimes" and 3 equaling "regularly"). At the end of the school year teachers' perception of students' frequency with which they demonstrated the selected skills increased on average by 0.22 to 2.81. The three items with the highest increase were the use of varied representations and approaches when solving problems, the maintaining oversight of the mathematical process while attending to detail, and the use of mathematical terminology and vocabulary with precision. The learning logs confirmed the theme of increased perseverance in problem solving which nine of twelve teachers as well as one coach reported. Furthermore, teachers reported students improved their ability to learn

cooperatively, student-to-student discourse increased, students were more engaged and more independent, and students improved their ability to explain their thinking. The coaches reported that students seemed more comfortable in their groups, and were better at working together productively, and exploring challenging concepts. The semi-structured interview responses confirmed those observations.

### *5. Teachers and students reported that instructional coaching professional development affected their use of teaching practices that implement the Common Core State Standards for Mathematical Practice.*

Program evaluation question five addressed the effect of the instructional coaching professional development as measured by the frequency of the use of teaching practices that evidence the implementation of the Common Core State Standards for Mathematical Practice. Instrument #7, in comparison with instrument #3, provided results to answer this question.

*Instrument #7: Student survey*

Five hundred and ninety-seven Algebra students (all Algebra students of the participating teachers) were given the opportunity to participate in an online survey at the end of the school year 2014-2015, asking them to rate the frequency of teaching practices they experienced in their Algebra classrooms (Guskey, 2000, level 4), and the frequency of their own learning coinciding with those strategies (Guskey, 2000, level 5), on a Likert scale 1 – 4, 1 equaling "never", 2 equaling sometimes, 3 equaling regularly, and 4 equaling "every lesson". The first part of the survey addressed the frequency of teaching practices they experienced in their Algebra classroom and contained 18 items. The

questions of the student survey mirrored the questions on the teacher survey regarding Guskey's (2000) levels but were worded from a student's perspective instead of the teachers'. Three hundred and fifty-nine students were able to take and submit the survey successfully.

The average of 359 students rated the frequency with which they experienced teaching practices that evidenced the implementation of Common Core Mathematical Practices (Appendix M) with 2.75 (2 equaling sometimes and 3 equaling regularly).

When comparing teachers perceived responses and the responses of their students as to with what frequency they experienced teaching strategies pertaining to the implementation of Common Core Mathematical Practices students rate the frequency on average lower than the teachers but confirm that those practices are occurring in their Algebra classrooms regularly (Appendix M).

***6. Teachers and students reported that instructional coaching professional development affected students' learning outcomes as measured by the demonstration of skills that evidence the implementation Common Core State Standards for Mathematical Practice.***

Program evaluation question six addressed the measurable effect of the instructional coaching professional development as measured by the frequency of the demonstration of skills that evidence the implementation Common Core State Standards for Mathematical Practice. Instruments #7 in comparison with instrument #3 provided results to answer this question.

*Instrument #7: Student survey*

The second part of the student survey asked the students to rate the frequency with which they demonstrated skills that evidenced the implementation of mathematical practices in their math classrooms. They were rating themselves on a Likert scale 1 – 4, 1 equaling "never", 2 equaling "sometimes", 3 equaling "regularly", and 4 equaling "every lesson". The second part contained 15 items.

The average of 359 students rated the frequency with which they perceived the demonstration of their mathematical skills that evidence the implementation of Common Core mathematical practices (Appendix M) with 2.71 (2 equaling sometimes and 3 equaling regularly).

When comparing teachers' responses and the responses of their students as to with what frequency they perceived students demonstrating mathematical skills evidencing the implementation of Common Core mathematical practices (Appendix M), students rated the frequency on average lower than the teachers but confirm that those skills were demonstrated in their Algebra classrooms almost regularly. Table 13 compared the teachers' pre-and post ratings regarding their knowledge and use of teaching practices as well as students' demonstration of skills that evidence the implementation of the Common Core State Standards for Mathematical Practice.

Table 13
*Comparison Between Mean Scores of the Different Levels of Guskey (2000)*

| Level | $\overline{x}$ Pre | $\overline{x}$ Post | Diff | Diff in % |
|---|---|---|---|---|
| **Level 2, teachers' knowledge, total average** | **2.94** | **3.25** | **0.31** | **10.5%** |
| **Level 4, use of knowledge and skills, total average** | **2.85** | **3.06** | **0.21** | **7.4%** |
| **Level 5, student learning outcomes, total average** | **2.59** | **2.81** | **0.22** | **8.5%** |

Level 2, teacher's knowledge, showed the highest increase, absolute and in percent. Level 4, transfer of knowledge and skills, showed the smallest increase, absolute and in percent. The total average post scores were highest for knowledge, followed by transfer of knowledge and skills, and smallest for student learning outcomes.

A paired t-test was conducted on pre and post survey scores to determine the statistical significance. Table 14 shows the survey items that resulted in a significant difference between pre and post mean scores.

Table 14
*Items with a Significant Difference Between Mean Scores of the Different Levels of Guskey*

| **Survey Item** | **N** | $\overline{x}$ **Pre** | $\overline{x}$ **Post** | **T** | **p** |
|---|---|---|---|---|---|
| **Level 2:** The asking of higher- level questions that require students to explain their thinking. | **12** | **3.0000** | **3.3333** | **-2.345** | **.039** |
| **Level 2:** The provision of multiple representations (models, number lines, tables, graphs, as well as symbols) to support visualization of skills and concepts. | **12** | **3.333** | **3.7500** | **-2.803** | **.017** |
| **Level 4:** I consistently ask higher- level questions that require students to explain their thinking. | **12** | **2.9167** | **3.2500** | **-2.345** | **.039** |

The results of the teacher pre and post survey as well as the student survey confirmed that instructional coaching professional development affected the use of teaching practices as well as the demonstration of skills that evidence the implementation of Common Core State Standards for Mathematical Practice positively. Two items of Guskey's level 2 resulted in a significant difference between the average pre and post test scores (as shown in Table 14) indicating that the instructional coaching professional development lead to an increased teacher knowledge about the asking of higher level questions that require students to explain their thinking, and the provision of multiple representations. One item of Guskey's level 4 resulted in a significant difference between the average pre and post test scores (as shown in Table 13) indicating that the

instructional coaching professional development lead to an to an increased teacher use of consistently asking higher level questions that require students to explain their thinking.

**Summary**

The results of the study demonstrated that instructional coaching had positive effects on teacher learning and teacher implementation of strategies as well as student learning outcomes. According to the responses in the interviews the context and experience of teachers with the instructional coaching professional development depended on a positive perception of teachers about the coaches and the possibility to learn new strategies. Teachers implemented strategies that were relevant to their classroom and their actual students. They appreciated the one on one work and the regular feedback to strategies they were implementing. It was important to consider teachers' requests (such as a model lesson) throughout the professional development to support them and create a positive environment for further visits. Teachers were willing to receive criticism but were sensitive how the feedback is delivered. If it was too critical and too many items were addressed they were discouraged and perceived the work with the coach as negative. Students learning outcomes did not manifest themselves necessarily in better assessment results even though some teachers made that connection. Student learning outcomes, according to the learning logs and interviews, were directly related to the strategies the teachers were implementing and it took time for the teachers to get comfortable and then for the students to successfully demonstrate new behaviors and learning.

## Chapter 5: CONCLUSIONS AND IMPLICATIONS

This final chapter includes a summary of the study and discusses the results of the program evaluation by addressing the overarching research questions. Conclusions and implications for evaluation practice and recommendations for future research then follow. The key findings of this study are discussed in light of their contribution to evaluation theory confirmation.

### Summary of the Study

The necessity to evaluate professional development has become more pressing for districts in a time of increased accountability. Research has identified features of effective professional development critical to ensure sustainable change in teaching practices and student learning: content focus, active learning, coherence, duration, and collective participation (Desimone, 2009, Joyce & Showers, 1995). Instructional coaching as a professional development model met these criteria (Knight, 2007). The district in this study chose an instructional coaching model to change teacher practice and improve student learning in mathematics and as part of common core implementation.

A number of empirical studies had shown that it was difficult to measure the effects of professional development, especially its effect on transferring practice and student achievement (Desimone, 2009; Garet et al., 2010; Guskey, 2000; Joyce & Showers, 2002; Reeves, 2010). Guskey and Yoon (2009) concluded that there were no reliable, valid, scientifically defensible data to show that alternative professional strategies such as instructional coaching worked. They challenged the field to critically assess and evaluate the effectiveness of professional development emphasizing the

necessity to meet a standard of rigor and the collection of meaningful and scientifically defensible data.

The purpose of this study was to use a realist approach to develop an evaluation based framework and to pilot the instruments to address the need as determined by the literature to effectively evaluate the instructional coaching professional development for mathematics.

This study was a mixed method study developing a framework to comprehensively evaluate the effectiveness of a coaching model. It was conducted at one comprehensive high school with the two instructional coaches, twelve Algebra teachers and their five hundred and ninety seven students as participants. The researcher used criterion sampling since the participating teachers had to receive instructional coaching professional development and work at the high school of this study. The selection of the research site was a convenient sample.

The researcher conducted a coach survey at the beginning of the year, a teacher pre-and post survey and a student survey at the end of the school year. During the school year teachers and coaches wrote learning log entries on a quarterly basis and at the end of the year the researcher conducted semi structured interviews with all the participating teachers. All data was able to address Guskey's (2000) five levels of professional development evaluation. Triangulation, member check, thick rich data (Creswell, 2009; Maxwell, 2014; Patton 2001) from interviews and the theory-based design (Maxwell & Mittapalli, 2010) were used to establish validity of the results.

## Discussion of the Results

The results of the evaluation questions will be discussed through a synthesis in light of a realist evaluation approach to address the overarching research questions.

### 1. *What 'context-mechanism-outcome pattern configurations' did the instructional coaching confirm in terms of the further refinement of the professional development?*

The purpose of a realist evaluation was to discover what it was about a program that worked for whom and in what circumstances. The concept was based on context, mechanism, and outcome patterns and ultimately their configurations. Context described the conditions, in which the program was introduced, that were relevant to the operation of the mechanisms. Mechanisms described what it was about a program that generated any effect. Outcome patterns included the intended and unintended consequences of programs due to the activation of different mechanisms in different contexts (Pawson &Tilley, 2004). According to Pawson and Tilley (2004) it used a configurational approach to causality and developed and tested the program theories that lead to assumed context-mechanism-outcome pattern configurations (CMOC's).

In order to synthesize and interpret the results of chapter 4 a realist hypothesis grid (see Table 15) was developed (Pawson & Tilley, 2004).

Table 15

*Context-Mechanism-Outcomes Configurations*

| Some Potential Contexts + | Some Plausible Mechanisms = | Some Possible Outcomes |
|---|---|---|
| Teachers participating in the pilot study had a positive perception of the coaching process and the coach | Coaches build a relationship and create trust | Teachers become open to change and want to learn different instructional practices |
| Teachers are open to change and want to learn different instructional practices | Coaches observe and give feedback to teachers | Teachers learn about different strategies |
| Teacher is not open to change and does not want to learn different instructional practices | Coaches observe and give feedback to teachers | Teachers do not learn about different strategies. |
| Coaches are positive and believe teachers have certain strengths to build on | Coaches encourage teachers in their areas of strengths | Teachers continue to use strategies that implement the mathematical practices |
| Teachers view themselves as always being able to get better in their craft | Coaches point out areas of potential improvement | Teachers reflect on their instructional practice and try new strategies |
| Teachers view criticism as evaluative and judgmental and undermining their professionalism | Coaches point out areas of potential improvement | Teachers feel criticized and do not change their practice |
| Teachers feel confident about new teaching strategy and safe to try them | Teachers try new strategies | Students respond to the new instructional practices in a positive way, change their behavior and demonstrate learning |
| Teachers do not feel confident about new teaching strategies and not safe to try them | Teachers do not try new strategies | Student behavior and learning does not change. |
| Teachers feel confident about new teaching strategies and safe to try them | Teachers try new strategies | Students do not respond to the new instructional practice, change their behavior, or demonstrate new learning |

*Note:* Adapted from "*The Science of Evaluation: A Realist Manifesto,"* by R. Pawson, p. 23. Copyright 2013 by R. Pawson.

The rationale for using the realist approach was that the evaluation was not seeking to determine if X caused Y but to consider the contextual mechanisms that were responsible for the relationship between X and Y. The table above lists some possible configurations that were aligned to the scope of this study.

The seven instruments measured some of these mechanisms and outcome patterns and/or provided explanations for them.

The configurations in Table 16 display how the professional development was expected to work based on the assumptions of Desimone's (2009) core action theory.

Table 16
*Context-Mechanism-Outcome Configurations Based on the Program Theory*

| **Contexts** | **+** | **Mechanisms** | **=** | **Outcomes** |
|---|---|---|---|---|
| Teacher is open to change and wants to learn different instructional practices | | Coaches observe and give feedback to teachers | | Teachers learn about different strategies |
| Coaches are positive and believe teachers have certain strengths to build on | | Coaches encourage teachers in their areas of strengths | | Teachers continue to use strategies that implement the mathematical practices |
| Teachers view themselves as always being able to get better in their craft | | Coaches point out areas of potential improvement | | Teachers reflect on their instructional practice and try new strategies |
| Teachers feel confident about a new teaching strategy and safe to try it | | Teachers try new strategies | | Students respond to the new instructional practices in a positive way, change their behavior and demonstrate learning |

*Note:* Adapted from "*The Science of Evaluation: A Realist Manifesto,"* by R. Pawson, p. 23. Copyright 2013 by R. Pawson.

*Outcomes*

In summary, teachers reported through the survey, learning logs and semi structured interviews that they perceived their knowledge about teaching practices that implement the mathematical practices of CCSS had increased through the instructional coaching professional development.

The teacher survey results (see Appendix K) showed an increase of the mean score of all teachers and all survey items of 10.5% for teachers' knowledge, 7.4% for use of knowledge and skills, and 8.5% for student learning outcomes. This suggested that

teachers knew more about certain instructional practices after the professional development, that they used certain strategies more regularly, and that they observed increased student learning outcomes connected to these strategies. Teachers reported that they perceived their use of teaching practices that implement mathematical practices of CCSS has increased through the instructional coaching professional development. The results of the student survey in comparison with the teacher survey (Appendix K) demonstrated that students as well as teachers perceived teaching practices and connected student learning outcomes to occur in the Algebra classrooms increased (.3 for use of knowledge and skills and .1 for student learning outcomes. Note that teachers rated both higher than the students). Teachers reported that the instructional coaching professional development had a positive effect on their students' learning defined as the demonstration of skills that evidence the implementation of the mathematical practices in the classroom. The results of the survey, learning logs, and interviews confirmed the above outcome patterns that teachers implemented new practices and students demonstrated their learning connected to those practices.

*Mechanisms*

Ten of twelve teachers explained that instructional coaching strengthened their implementation of teaching practices, which was due to the one on one approach with immediate feedback and the relevance to their classroom. Four teachers confirmed that suggestions by the coaches and their reviews of the classes was supportive. They were given specific ways in which to modify particular lessons to provide students more clarity. All teachers reported that they would like to continue instructional coaching professional development and made suggestions for improvement. Ten teachers

perceived the professional development as very beneficial, teacher specific, and relevant to their teaching. The coaches' survey asked the participating teachers to rate 22 items that coaches were responsible for doing, 21 items were rated between at least sometimes and regularly. The items with the lowest scores support the results from the semi structured interviews that teachers felt the coaches did not support the planning of the lessons but were more reactive due to the scheduling and the design of the session with the coaches (feedback after a lesson). The highest ratings suggest that the coaches were regularly meeting with the teachers, provided feedback, and set high standards for teacher performance. The teachers reported that the coaches regularly contributed positively to the improvement of instruction and that their instructional coach was a positive person who maintained his/her composure. The coaches' survey and semi structured interview results confirmed the above mechanisms took place regularly for a majority of the teachers. One teacher made the following statement: "You know that's been the biggest thing, I feel respected and like a partner with the coach as opposed to an underling." Another teacher described his/her experience with the coach as follows: The instructional coaching was more one on one and my instructional coach was able to help me with what my weaknesses were, what my needs were individually. So he/she was able to adapt to what I needed which was great".

*Context*

The context of the professional development was described in the semi structured interview results. Prior to the instructional coaching professional development, teachers went to an off site professional development session a few times a year and received a

lecture style presentation on some topic. Most teachers reported that they never implemented any of the strategies but perceived the gathering with all colleagues from the district as beneficial. The Algebra 1 teachers, three of which participated again this year, piloted the coaching professional development the previous year. They indicated a positive perception of working with their coach and shared the results with the rest of the staff. The cultural context for the professional development was one of willingness to try something new and a sense of support, safety, and professional respect from the coach. All teachers perceived themselves as willing to improve their practice and to receive constructive suggestions.

*Unintended context-mechanisms outcome configurations*

Even though the results evidenced some configurations as intended by the professional development, the realist evaluation examined possible configurations that were not intended and what occurred that lead to undesired outcomes.

Table 17 shows some of the possible configurations that were not intended by the program.

Table 17
*Unintended Context-Mechanism-Outcome Configurations*

| Contexts | + | Mechanisms | = | Outcomes |
|---|---|---|---|---|
| Teacher is not open to change and does not really want to learn different instructional practices | | Coaches observe and give feedback to teachers | | Teachers do not learn about different strategies. |
| Teachers view criticism as evaluative and judgmental and undermining their professionalism | | Coaches point out areas of potential improvement | | Teachers feel criticized and do not change their practice |
| Teachers do not feel confident about a new teaching strategy and not safe to try it | | Teachers do not try new strategies | | Student behavior and learning does not change. |
| Teachers feel confident about a new teaching strategy and safe to try it | | Teachers try new strategies | | Students do not respond to the new instructional practice, change their behavior, and demonstrate learning |

*Note:* Adapted from "*The Science of Evaluation: A Realist Manifesto,"* by R. Pawson, p. 23. Copyright 2013 by R. Pawson.

It was more difficult to explain these configurations and find confirmatory evidence. The survey data indicated that teachers did not rate themselves on average as having extensive knowledge after the professional development or using all the teaching practices every lesson. Student survey results reported that students perceived teachers to use any practice between sometimes and regularly, so there appears to be potential for future growth.

The results of the coaches' survey per teacher showed one teacher rating the performance of the coach significantly lower than the other participants which indicated that the mechanisms did not take place as intended. The semi-structured interviews provided more detailed explanations regarding unintended mechanisms. Two teachers felt the instructional coaching professional development was not beneficial. One explained his/her lack of learning was due to the quality of feedback that was received. The teacher also felt that peer coaching would be more beneficial than an outside expert coming in.

but also noted that it could have been the particular personality of the coach that did not work as well as the other coach who seemed to work better. The other teacher explained the lack of learning was due to critical early feedback at the beginning of the school year that made him/her feel as if she did everything wrong. Both teachers noted that the coach took their feedback and suggestions and the conversations became more encouraging. Additionally, the coach was willing to teach a model lesson on request and it was later mentioned by most teachers that this had been very helpful to all of them. Even though there was no evidence that suggested teachers were resistant to instructional coaching, the two teachers were not able to learn much from the coach.

Student learning outcomes were reported as observed by the teachers and as reported by the students. All teachers that felt the professional development was beneficial indicated that they saw some student learning as summarized before. However teachers also noted that only certain students were receptive to the new practices while others did not engage during class or do their homework. One teacher specifically stated that students were more engaged in group work and seemed to grasp the concepts better but then did not perform well in assessments during the first quarter. Later in the school year, the teacher could see improvement in students' conceptual understanding evidenced by the retention of knowledge by the time of the final exam review.

*Recommendations for change of the Professional development*

The pilot study used teacher suggestions to broaden the focus of the professional development, add new challenges, and to consider peer coaching with a different person to gain a new perspective. Additionally, teachers expressed an interest in having pre-

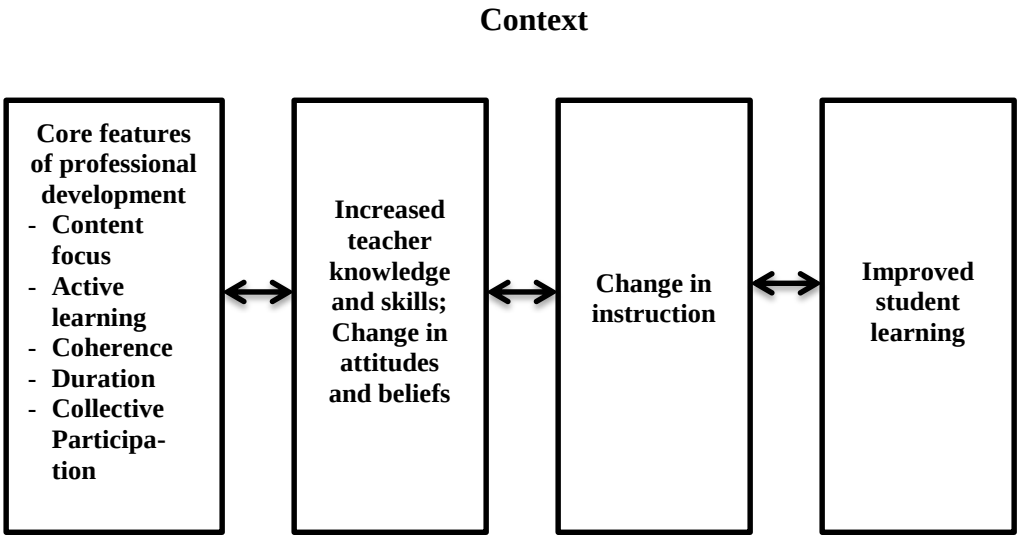
observation conferences instead of post observation conferences and to schedule less frequent visits, but more at the beginning of the school year to have time to implement strategies. Regarding the content of the professional development, teachers mentioned a desire to learn more content specific strategies, how to teach a particular concept, and to provide the teachers with specific resources such as technology and manipulatives.

### *2. Was it possible to create a 'realist evaluation' based framework to evaluate an instructional coaching model of professional development?*

Guskey and Yoon (2009) summarized the research connecting the effects of professional development to student learning as sparse when only 9 out of 1343 quantitative studies met the standards of credible evidence set by the What Works Clearinghouse. "The amount of valid and scientifically defensible evidence we currently have on the relationship between professional development and improvements in student learning is exceptionally modest" (p.499). According to the National Mathematics Advisory Panel's report of 2008, most studies of professional development in mathematics were descriptive and lacked methodological rigor that was needed to warrant causal inferences (e.g. one group pretest/posttest designs without a comparison group, in Guskey & Yoon, 2009). They charged the planners of professional development to consider not only the goals but also what evidence best reflected the achievement of those goals, and how the evidence could be gathered in meaningful and scientifically defensible ways (Guskey & Yoon, 2009). Lastly they specifically addressed the advocates of alternative professional development models to "take responsibility for demonstrating effectiveness through rigorous and scientifically valid means" and to "take

the time to conduct thorough and systematic investigations of the true effects" (p.498). Cornett and Knight (2009) after summarizing existing research on the effects of instructional coaching state that the field needed further research since much of the research had been exploratory and lacked the rigor of a true scientific study. Various coaching models were in their early stages and program designers were interested in feedback from participants and quick data gathering. Desimone (2009) suggested a conceptual framework for studying the effects of professional development on teachers and students (see Figure 8).

Figure 8. Conceptual Framework for the Effects of Professional Development on Teachers and Students(such as teacher and student characteristics, curriculum, school leadership, policy environment)




*Note:* Adapted from "Improving impact studies of teacher's professional development: Toward better conceptualizations and measures", by L. M. Desimone, 2009, *Educational Researcher, 38*(3), p. 185. Copyright 2009 by AERA.

Desimone (2009) stated that context was an important mediator and moderator and listed several key features derived from the literature such as student and teacher characteristics, contextual factors in the classroom, school, and district level, and policy conditions. This study sought to address the requirements stated in the literature while

applying the frameworks to a local setting where practicality, time constraints, and limited resources had to be considered. Professional development decisions in an average size district with three high schools are not made based on large scale policy studies and the methods to evaluate the professional development to inform the district and building leadership have to be appropriate for the setting. Guskey & Yoon (2009) state that experimental or quasi-experimental studies provide the most valid and scientifically defensible evidence available, but such approaches are difficult in a practical local setting. If an intervention or particular professional development is to be considered effective at improving student learning, students cannot be excluded for the purpose to create a control group. The district or building administration however would still like to know what the effects of the professional development were, if they should continue to offer it, and what improvements should be made. Furthermore, the literature suggested that the number of studies that were able to provide meaningful results were sparse. Quasi-experimental designs still require the isolation of a testable variable while all other conditions are controlled for. Alternative professional development models such as instructional coaching were even more contextual and further challenge evaluation designs that were able to produce valid causal inferences.

This study sought to meet Guskey's and Yoon's (2009) request to demonstrate effectiveness "through rigorous and scientifically valid means" and to conduct a "thorough and systematic investigation of the true effects" (p.498). However, terms such as "effectiveness", and "scientifically valid means" depend on the research paradigm one subscribes to. This researcher wanted to investigate if conducting an evaluation using a realist approach is an effective means of evaluating instructional coaching. The rationale

for this approach lies in the realist interpretation that causality refers to actual causal mechanism and processes that were involved in particular events and situations (Maxwell & Mittapalli, 2013). Pawson and Tilley (1997) stated that "the idea is that the mechanism is responsible for the relationship itself. A mechanism is ….not a variable but an account of the make up, behavior, and inter-relationship of those processes which are responsible for the regularity" (p.67-68).

This study based its evaluation on Desimone's (2009) core action theory and Guskey's (2000) levels of evaluation, but intended to explain and measure outcome patterns as a result of mechanisms that the theory suggested as plausible to happen instead of attempting to isolate variables and measure their effect. Due to the complex system in which instructional coaching takes place, and the many variables that influence the occurrence of coaching, a process view of causality was the appropriate approach to develop the instruments.

The instruments were able to provide evidence of mechanisms taking place and generating outcome patterns such as increased teacher knowledge, increased teacher learning, and increased student learning. The instruments measured these outcomes by tracing causal paths. Figure 9 for example shows the assumed causal path on the teacher survey for one particular strategy.

Figure 9. Causal Path of One Teaching Practice on the Teacher Survey

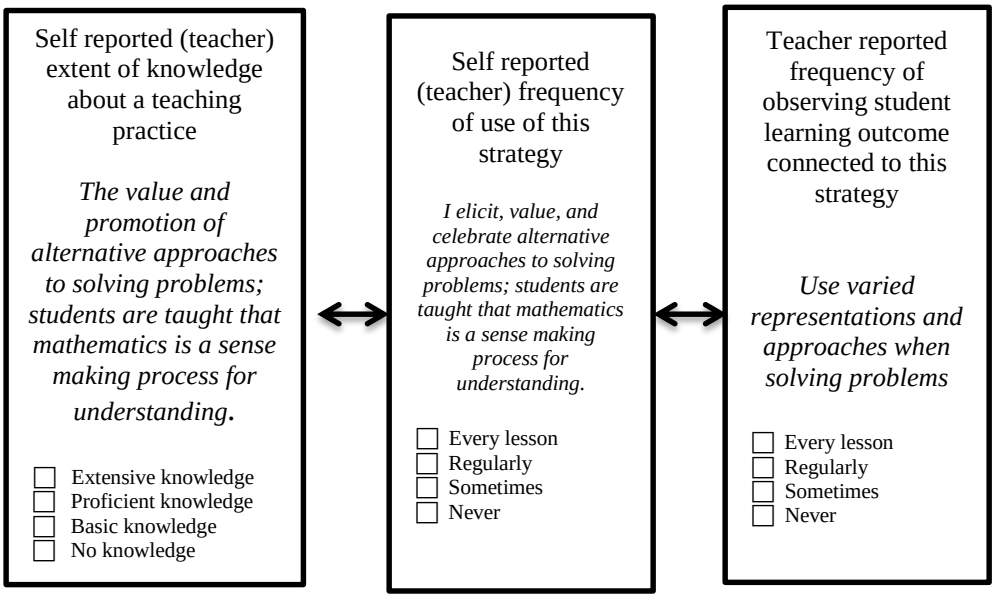
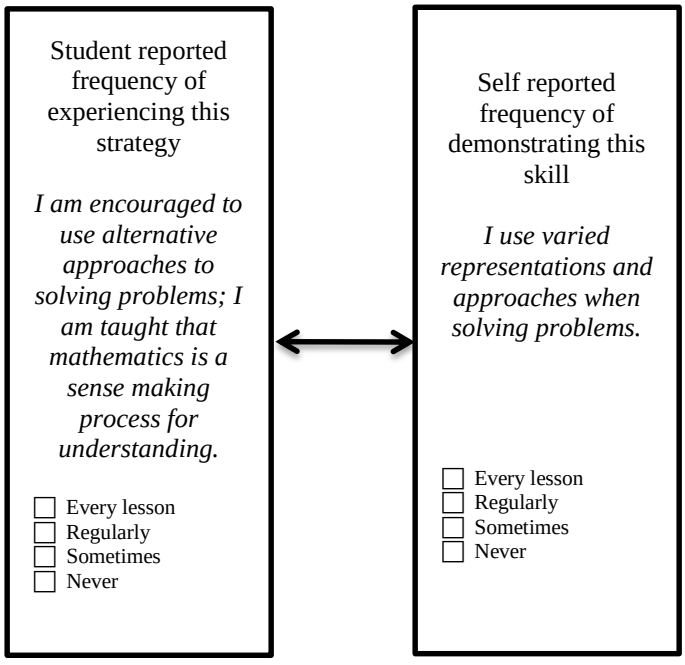



Figure 10 shows the causal path for the same teaching practice on the student survey (the use of varied representations and approaches when solving problems).

Figure 10. Causal Path of One Teaching Practice on the Student Survey



The construction of the surveys allowed for the collection of mean score measures of the level of knowledge, implementation, and aligned student learning outcomes as a result of mechanisms working according to intended professional development design as perceived by teachers and students. The coach survey provided evidence of mechanisms and outcome patterns regarding the work of the coaches. Learning log and semi-structured interviews provided explanations to what worked for whom and what did not work and why.

In summary, the design of the instruments allowed for an evaluation of the instructional coaching professional development following the guidelines and paradigm of a realist approach to evaluation. The construction of the surveys applied a process view of causality to measure the outcome patterns that would occur through plausible mechanisms according to the program theory (Pawson and Tilley, 2004). The learning logs and surveys added qualitative data to provide causal explanations for the context in which the professional development took place and were able to inform program refinements for the next school year (Maxwell, 2012).

## Conclusions

In conclusion, the realist approach to evaluate instructional coaching professional development is an appropriate design to consider due to the contextual nature of coaching. The program theory to be confirmed must demonstrate plausible context-mechanism-outcome pattern configurations rather than simplified ad hoc explanations (Pawson, 2013). The questions of ‘what works for whom and when and in what respect’

seems most appropriate for the program evaluation of professional development that a district is seeking to evaluate and refine.

The construction of the instruments based on established professional development evaluation models such as Guskey's (2000) ensures a sound theory base that increases the credibility and validity of the results. Furthermore, the construction of the surveys to link plausible mechanisms and plausible outcomes following the logic of the program theory allowed the results to convey important detailed information regarding different components that the professional development originally targeted.

The inclusion of qualitative data such as the learning logs and the semi-structured interviews added the necessary explanations and context to the quantitative data and was essential to a realist approach that sought to answer how and why questions.

However, the complexity of using the realist approach to frame a program evaluation remains a challenge (Pawson 2013). While the plausible context- mechanism-outcome pattern configurations were theory based, there are infinite alternative scenarios that could also hold true. The theory provides the assumptions for planners of how and why interventions might work (Pawson, 2013).

This study provided the district leaders and building administration a better understanding of what aspects of the professional development should be continued, what the strengths and weaknesses were, and what could be changed to address the concerns of the teachers.

## Implications

The implications of this study are threefold: First, any district and school that conducts professional development should identify and consider the program's theory. Before making decisions regarding the evaluation design for the professional development, the intended causal paths of the professional development should be clarified. The underlying theory provides the assumptions for causal connections between activities and particular outcomes (Bichelmeyer, & Horvitz, 2006; Rogers et al., 2000). This provides better evidence for causal attribution (Rogers et al., 2000). The program theory can be visualized and described with a logic model that links short-term and long-term outcomes with the activities and processes. The evaluation needs to systemically investigate the defined key performance measurement points by its sequences (Cooksy, Gill, & Kelly, 2001; McLaughlin & Jordan, 2010). Logic models additionally facilitate a way to communicate to stakeholders in the planning stage of the professional development that informs the evaluation. Without a theory based approach, causality in contextual settings becomes "ad hoc" (Pawson, 2013, p.27), and the evaluation has no sound foundation to stand on in regards to what to measure and why.

Second, evaluators need to design the evaluation to match the program theory considering contextual mechanisms. Practical evaluations from school districts such as Friendship Charter School (2011), Ithaca City School District (2013), and Erie Public Schools (GE Foundation, 2013) all employed multiple instruments, mostly surveys, interviews and some type of protocol to observe the frequency of implementation. Student achievement was not addressed. The evaluations used various methods but did not base them on an explicit program theory. Even though the data provided useful

information, it could not speak to the mechanisms and how they worked and under what conditions they generated the outcomes that were measured. The data collection decisions did not follow an evident logic derived from a program theory. This warrants the criticism from Guskey and Yoon (2009) that practitioners need to demand better evidence in regards to the effectiveness of new strategies and practices and more rigorous and meaningful evaluations need to be conducted to show true effects of a professional development. The choice of data collection has to be tailored to the hypothesis or assumptions about what mechanisms generate what outcomes (McLaughlin & Jordan, 2010; Pawson & Tilley, 1997).

Third evaluators need to be clear when designing an evaluation of what kind of results the evaluation can and cannot provide due to the design and what questions can be answered by the evaluation. These choices and resulting limitations need to be clearly communicated to stakeholders. The more contextual the nature of the professional development, the less likely the evaluation can answer the question "did it work" with scientific, valid measures. Newcomer, Hatry and Wholey (2010) pointed out that it was very difficult to draw causal inferences and to capture the net impact of a program. They suggested to use terms such as 'plausible attribution' when drawing conclusions about the effects of a program. If the collected data is more qualitative and is answering why and how questions, stakeholders need to be aware of what kind of decisions the results can inform and what limitations these causal explanations contain.

**Recommendations for Future Research**

Research evaluating contextual professional development designs such as instructional coaching are beginning to be developed. Cornett and Knight (2009) ask evaluators to address all of Guskey's (2000) levels when evaluating coaching. This study provided the application of an evaluation approach that matched the contextual nature of the professional development design that it sought to evaluate and based the framework on Guskey's (2000) five levels of evaluation. However, there is a need for future research to apply a 'realist evaluation' approach to professional development evaluation utilizing different methods and instruments to test context-mechanism-outcome configurations (CMOC's). "If a cardinal purpose of evaluations is to feed into improvements in policy and practice, they need to be oriented to culmination" (Pawson &Tilley, 1997; p. 115). The analysis of CMOC's of prior studies can build and inform further investigations of professional development to continue to develop theory. As studies propose and test CMOC's further, studies can test the same CMOC's in different settings to refine the professional development and gain insight under what conditions intended outcomes occur. Cumulative evaluations using a realist approach can meet Guskey and Yoon's (2009) request for trustworthy, verifiable, replicable, and comparative data to increase credibility and acceptance for alternative professional development designs. Furthermore they can contribute to increased professionalism in the field of professional development evaluation.

## Summary

This study demonstrated the application of a 'realist' approach to evaluate instructional coaching professional development. The framework was developed to test and explain context-mechanism outcome configurations that the program theory suggested. The findings addressed all five levels of Guskey's (2000) evaluation model and indicated an increase in teacher knowledge (level 2), teacher use of knowledge and skills (level 4) and student learning outcomes (level 5). The results further described the context and mechanisms that lead to these outcomes, namely relevant one on one professional development within a safe, respectful professional environment that fostered an interest and willingness to change practice.

## Appendix A
## Consent forms participants

### Consent form teachers and coaches

**Title of research study:**
Instructional Coaching Evaluation Using a Realist Approach: A Mixed Methods Study

**Investigator:** Claudia Berlage
We invite you to take part in a research study because you are either the direct recipient of the professional development (Algebra teacher) or are administering the professional development (consultant, Teacher On Special Assignment)

**What you should know about a research study**

- Someone will explain this research study to you.
- You volunteer to be in a research study.
- Whether or not you take part is up to you.
- You can choose not to take part in the research study.
- You can agree to take part now and later change your mind.
- Whatever you decide it will not be held against you.
- Feel free to ask all the questions you want before you decide.

**Who can I talk to?**
If you have questions, concerns, or complaints, or think the research has hurt you, talk to the research team at the University of Bridgeport.

This research has been reviewed and approved by an Institutional Review Board. You may talk to the IRB Co-Chair at (203) 576-4141 or irb@bridgeport.edu about any of the following:

- Your questions, concerns, or complaints are not being answered by the research team.
- You cannot reach the research team.
- You want to talk to someone besides the research team.
- You have questions about your rights as a research subject.
- You want to get information or provide input about this research.

**Why are you doing this research?**
The purpose of this study is to gain insight to evaluate the classroom embedded professional development in Mathematics to inform the design of future PD as well as the development of a framework to evaluate classroom embedded PD that effectively supports you in meeting your students' needs.

**How long will the research last?**
We expect that you will be in this research study for 10 months till the end of the school year 2014/2015.

**How many people will be studied?**

19 teachers, 3 coaches, 2 TOSA's

**What happens if I say yes, I want to be in this research?**

You will participate in a 30-45 minute interview with semi-structured interview questions asking you to share your experiences with this professional development at the end of the school year 2014-2015. The interview will be held by the investigator. You will fill out a formative survey at the beginning of the school year and the same survey as a summative survey at the end of the school year (which should take about 20 minutes). You will describe your own learning one time per quarter with a learning log entry (approximately 30 minutes for each entry). You will also fill out a survey regarding the instructional coach (10 Minutes) at the end of the school year 2014/2015.

**What happens if I say no, I do not want to be in this research?**

You may decide not to take part in the research and it will not be held against you.

**What happens if I say yes, but I change my mind later?**

You agree to take part in the research now. You may stop at any time and it will not be held against you.

**Is there any way being in this study could be bad for me?**

***N/A***

**Will being in this study help me any way?**

We cannot promise any benefits to you or others from your taking part in this research. However, possible benefits include your insight shaping and informing future professional development, improving instruction and learning.

**Is there any way being in this study could be bad for me?**

***N/A***

**What happens to the information you collect?**

Your information will be confidential. Results will be shared with the district but the identity of the participants kept confidential. The surveys confidential and the identity of the author of the learning logs will be confidential.

We may publish the results of this research. However, we will keep your name and other identifying information confidential.

**Can I be removed from the research without my OK?**
N/A

| | |
|---|---|
| Signature of subject | Date |
| Printed name of subject | |
| Signature of person obtaining consent | |

**Passive Consent Parents**

Date
Dear Parent/Guardian:
We are writing to tell you about a research project that is being conducted at X High School as part of a doctorate program.
The purpose of this research project is to evaluate the classroom embedded professional development in Mathematics at X High School during the school year 2014-2015. The district wishes to evaluate the effectiveness of the professional development in Mathematics and its effect on student learning.
This is a research project being conducted by Claudia Berlage (also Assistant Principal at X High School), at the University of Bridgeport. Your child is invited to participate in this research project because she/he is a Algebra student at Stamford High School.

The procedure involves filling out an online survey that will take approximately 20 minutes. Your child's responses will be confidential and we do not collect identifying information such as your child's name, email address or IP address. The survey questions will be about the mathematics instruction in your child's Algebra classroom.

We will do our best to keep your information confidential. All data is stored in a password protected electronic format. To help protect your confidentiality, the surveys will not contain information that will personally identify you. The results of this study will be used for scholarly purposes only and may be shared with University of Bridgeport representatives and once the research is published, with the district.

If you have any questions about the research study, please contact Claudia Berlage (cberlage@bridgeport.edu). This research has been reviewed according to University of Bridgeport IRB procedures for research involving human subjects.

Your child's participation in this research study is voluntary. You may choose not to participate. If you decide for your child not to participate in this study your child will not be penalized.

We are contacting you to ensure that you know that if you DO NOT want your child to participate that you may opt out by completing the opt-out online form or by returning the opt-out form on the back of this letter and sending to the attention of:

Chief Information Officer

All opt-out forms must be returned by 05/15/2015 as this phase of the study is scheduled to begin on 05/30/2015. If you have any questions, please feel free to contact Claudia Berlage at 203 977 4268.

Sincerely,
Chief Information Officer

Claudia Berlage
Assistant Principal
Doctoral Candidate
University of Bridgeport

## 2015 Research Study Survey
## OPT-OUT FORM

Complete the information below only if you do not wish your child to participate in the Evaluation of Mathematics Professional Development Research study. Please make sure that any opt-out forms are completed prior to 05/15/2015

---

☐ I do not give permission for my child to participate in the 2015 Evaluation of Mathematics Professional Development research study being conducted at X High School. I understand that my child will be assigned to a different activity during the administration of the survey.

Student's Name: _____
Student's Grade:_____
Student's Teacher:_____
Signature of Parent/Guardian: _____

Please return this page to your student's school office only if you ***do not wish your student to participate.*** It will be assumed that your child has permission to participate in the research study survey, if we do not complete this form by 05/15/2015. We appreciate your assistance and cooperation.

## Appendix B
## Instrument # 1: Instructional Coach survey for faculty

(Adapted and abbreviated from Instructional Coach Faculty Survey Alabama PEPE program)

Please respond to all statements by circling one number or letter to the right of the statement. Respond from your own knowledge about this instructional specialist using the following rating scale with the corresponding response choices: 1 - rarely, 2 - sometimes, 3 - usually, 4 - almost always, x - don't know

| **Statement** | Rarely | Sometime | Usually | Almost Always | Don't Know |
|---|---|---|---|---|---|
| 1. This instructional coach meets with me during the scheduled time. | 1 | 2 | 3 | 4 | x |
| 2. This instructional coach helps me overcome barriers to teaching and learning. | 1 | 2 | 3 | 4 | x |
| 3. This instructional coach helps me evaluate my instruction. | 1 | 2 | 3 | 4 | x |
| 4. This instructional coach helps me to identify and solve problems. | 1 | 2 | 3 | 4 | x |
| 5. This instructional coach finds new ways to do things better. | 1 | 2 | 3 | 4 | x |
| 6. This instructional coach contributes positively to improvement of instruction. | 1 | 2 | 3 | 4 | x |
| 7. This instructional coach believes in celebrating instructional and academic improvement. | 1 | 2 | 3 | 4 | x |
| 8. This instructional coach maintains open, two-way communication with school faculty**.** | 1 | 2 | 3 | 4 | x |
| 9. This instructional coach sets high standards for teacher performance. | 1 | 2 | 3 | 4 | x |
| 10. This instructional coach performs duties in a professional manner | 1 | 2 | 3 | 4 | x |
| 11. This instructional coach communicates information clearly and succinctly. | 1 | 2 | 3 | 4 | x |
| 12. This instructional coach communicates the importance of focusing on the needs of students. | 1 | 2 | 3 | 4 | x |

| **Statement** | Rarely | Sometime | Usually | Almost Always | Don't Know |
|---|---|---|---|---|---|
| 13. This instructional coach promotes and supports innovations. | 1 | 2 | 3 | 4 | x |
| 14. This instructional coach is a positive person who maintains his/her composure. | 1 | 2 | 3 | 4 | x |
| 15. This instructional coach helps me establish routines and procedures that contribute to learning and teaching of Math content. | 1 | 2 | 3 | 4 | x |
| 16. This instructional coach conducts planning and feedback sessions with me. | 1 | 2 | 3 | 4 | x |
| 17. This instructional coach works with me to ensure that standards and instruction are aligned. | 1 | 2 | 3 | 4 | x |
| | 1 | 2 | 3 | 4 | x |
| 18. This instructional coach provides feedback to me about my instruction. | 1 | 2 | 3 | 4 | x |
| 19. This instructional coach assists me in instructional planning when needed. | 1 | 2 | 3 | 4 | x |
| 20. This instructional coach understands research-based instructional procedures and helps me implement these procedures. | 1 | 2 | 3 | 4 | x |
| 21. This instructional specialist keeps abreast of teaching/learning research and best practices. | 1 | 2 | 3 | 4 | x |
| 22. This instructional coach is a mentor to me. | 1 | 2 | 3 | 4 | x |

(Adapted and abbreviated from Instructional Coach Faculty Survey Alabama PEPE program)

## Appendix C
## Instrument # 2: Semi-structured interviews with teachers
## (Researcher developed protocol)

The following interview will be taped. Before you are the semi-structured interview questions.

**Semi-structured interview questions:**

*Context*

1.) How has district Math professional development been designed prior to this instructional coaching professional development?
2.) What were the strengths and weaknesses of that professional development

*Guskey level 1: Teachers' reaction*

3.) How was the instructional coaching professional development different?
4.) What about the instructional coaching professional development would you continue to do and why?
5.) What would you change and why?

*Guskey level 2: Teachers' learning*

6.) What have you learned through instructional coaching professional development that makes you a more effective teacher?
7.) What would you like to explore further in a professional development next year?
8.) In what way did the instructional coaching professional development change your perceptions of your teaching practice?

*Guskey level 3: Organization support*

9.) How has your collaborative team (IDT team or cluster) positively or negatively influenced the professional development?
10.) How has scheduling influenced the success or challenge with the professional development?

*Guskey level 4: Teachers' use of new knowledge and skills*

11.) What instructional strategies have you implemented on a regular basis as a result of the embedded professional development?
12.) What do you think about the instructional coaching professional development as a way to strengthen implementing CCSS?

*Guskey level 5: Students' learning outcomes*

13) How do you think has the instructional coaching professional development affected student learning?

## Appendix D

**Instrument #3: Formative and summative teacher survey**
**(Researcher developed instrument)**

Today's date:
Course: Algebra 1 ☐ Algebra 2 ☐
Pilot teacher: Yes ☐ No ☐
Name of teacher:
Years of experience teaching this course:

**Teacher's learning**
**(Guskey level 2)**
(from GE CCSS Math classroom observation guide)

**Describe your knowledge level regarding the following items:**
(Extensive - 4, Proficient - 3, Basic -2 , No knowledge -1)

The instructional shifts in mathematics.
☐ Extensive knowledge
☐ Proficient knowledge
☐ Basic knowledge
☐ No knowledge

The maintenance of a high cognitive demand throughout the lessons, requiring students to deeply engage with making sense of the mathematics and justifying their thinking.
☐ Extensive knowledge
☐ Proficient knowledge
☐ Basic knowledge
☐ No knowledge

The asking of higher- level questions that require students to explain their thinking.
☐ Extensive knowledge
☐ Proficient knowledge
☐ Basic knowledge
☐ No knowledge

The review of critical prerequisite skills and concepts.
☐ Extensive knowledge
☐ Proficient knowledge
☐ Basic knowledge
☐ No knowledge

The application of math concepts to real world situations.
- ☐ Extensive knowledge
- ☐ Proficient knowledge
- ☐ Basic knowledge
- ☐ No knowledge

The eight Mathematical Practices
- ☐ Extensive knowledge
- ☐ Proficient knowledge
- ☐ Basic knowledge
- ☐ No knowledge

The development of number sense by asking for estimates, mental calculations, and equivalent forms of numbers.
- ☐ Extensive knowledge
- ☐ Proficient knowledge
- ☐ Basic knowledge
- ☐ No knowledge

The value and promotion of alternative approaches to solving problems; students are taught that mathematics is a sense making process for understanding.
- ☐ Extensive knowledge
- ☐ Proficient knowledge
- ☐ Basic knowledge
- ☐ No knowledge

The provision of multiple representations (models, number lines, tables, graphs, as well as symbols) to support visualization of skills and concepts.
- ☐ Extensive knowledge
- ☐ Proficient knowledge
- ☐ Basic knowledge
- ☐ No knowledge

The use of mathematical terminology and vocabulary with precision.
- ☐ Extensive knowledge
- ☐ Proficient knowledge
- ☐ Basic knowledge
- ☐ No knowledge

Class routines that are working effectively to facilitate learning.
- ☐ Extensive knowledge
- ☐ Proficient knowledge
- ☐ Basic knowledge
- ☐ No knowledge

Students' active engagement in lesson activities. Students are on task even when working independently.

- [ ] Extensive knowledge
- [ ] Proficient knowledge
- [ ] Basic knowledge
- [ ] No knowledge

The provision of opportunities for students to engage with the central content within all aspects of the lesson

- [ ] Extensive knowledge
- [ ] Proficient knowledge
- [ ] Basic knowledge
- [ ] No knowledge

Students' engagement in productive classroom discussions making connections and collaborating with others.

- [ ] Extensive knowledge
- [ ] Proficient knowledge
- [ ] Basic knowledge
- [ ] No knowledge

The careful planning of tasks, activities, questions, and assessments for coherence.

- [ ] Extensive knowledge
- [ ] Proficient knowledge
- [ ] Basic knowledge
- [ ] No knowledge

Instruction's orientation toward multiple levels of difficulty. All students are working at an appropriate level.

- [ ] Extensive knowledge
- [ ] Proficient knowledge
- [ ] Basic knowledge
- [ ] No knowledge

The availability of multiple activities and choices for student learning. A comprehensive set of strategies is used.

- [ ] Extensive knowledge
- [ ] Proficient knowledge
- [ ] Basic knowledge
- [ ] No knowledge

Students' choice how they will demonstrate their learning.
- ☐ Extensive knowledge
- ☐ Proficient knowledge
- ☐ Basic knowledge
- ☐ No knowledge

The use of formative assessment to consistently check on lesson objectives/targets.
- ☐ Extensive knowledge
- ☐ Proficient knowledge
- ☐ Basic knowledge
- ☐ No knowledge

**Transfer of knowledge and skills**
**(Guskey level 4)**
(From Instructional Implementation Sequence: Attaining the CCSS Mathematical Practices Engagement Strategies and GE CCSS Math classroom observation guide)

**How frequently do you engage in the following teaching strategies?**
(Every lesson - 4, Regularly - 3, Sometimes - 2, Never - 1)

I initiate think, pair share
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

I require students' to show (communicate) their thinking in classrooms
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

I question and wait
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

I use grouping and engaging problems
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

I use questions and prompts with groups

- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

I allow students to struggle and encourage them to persevere

- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

I encourage students to explain their reasoning

- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

I maintain high cognitive demand throughout my lessons, requiring students to deeply engage with making sense of the mathematics and justifying their thinking.

- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

I consistently ask higher- level questions that require students to explain their thinking.

- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

I begin my lessons with a review of critical prerequisite skills and concepts.

- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

My students and I co-construct the application of math concepts to real world situations.

- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

I take every opportunity to develop number sense by ask for estimates, mental calculations, and equivalent forms of numbers.

☐ Every lesson
☐ Regularly
☐ Sometimes
☐ Never

I elicit, value, and celebrate alternative approaches to solving problems; students are taught that mathematics is a sense making process for understanding.
☐ Every lesson
☐ Regularly
☐ Sometimes
☐ Never

I provide multiple representations (models, number lines, tables, graphs, as well as symbols) to support visualization of skills and concepts.
☐ Every lesson
☐ Regularly
☐ Sometimes
☐ Never

I use mathematical terminology and vocabulary with precision.
☐ Every lesson
☐ Regularly
☐ Sometimes
☐ Never

My class routines are working effectively to facilitate learning.
☐ Every lesson
☐ Regularly
☐ Sometimes
☐ Never

All my students are actively engaged in lesson activities. Students are on task even when working independently.
☐ Every lesson
☐ Regularly
☐ Sometimes
☐ Never

All aspects of my lesson provide opportunities for students to engage with the central content.

- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

My students engage in productive classroom discussions making connections and collaborating with others.

- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

I have carefully planned tasks, activities, questions, and assessments for coherence.

- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

My instruction is oriented toward multiple levels of difficulty. All students are working at an appropriate level.

- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

I make multiple activities and choices available for student learning. A comprehensive set of strategies is used.

- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

I allow students to choose how they will demonstrate their learning.

- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

I use formative assessment to consistently to check on lesson objectives/targets.

- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

**Student learning**
**(Guskey level 5)**
(from GE CCSS Math classroom observation guide)

**I see the majority of my students do this during class:**
(Every lesson - 4, Regularly - 3, Sometimes - 2, Never - 1)

**1. Make sense of problems and persevere in solving them**
Persevere in solving difficult and worthwhile problems
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

**2. Reason abstractly and quantitatively.**
Make sense of quantities and their relationships in problem situations
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

Use varied representations and approaches when solving problems
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

Know and flexibly use different properties of operations and objects
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

**3. Construct viable arguments and critique the reasoning of others.**
Explain their thinking
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

Build upon their own and others' thinking
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

Critique the arguments and reasoning of others
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

**4. Model with mathematics.**
Apply the mathematics they know to solve problems arising in everyday life and the workplace
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

Analyze mathematical relationships to draw conclusions
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

Apply what they know and are comfortable making assumptions and approximations
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

**5. Use appropriate tools strategically.**
Consider the available tools when solving a mathematical problem
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

**6. Attend to precision.**
Use mathematical terminology and vocabulary with precision
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

**7. Look for and make use of structure.**
Look for, develop, generalize, and describe a pattern orally, symbolically, graphically, and in written form
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

**8. Look for and make use of regularity in repeated reasoning.**
Notice if calculations are repeated and look both for general methods and for short cuts
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

Maintain oversight of the mathematical process while attending to detail.
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

**59 questions**

## Appendix E

**Instrument #4: Teacher learning logs**
**(Researcher developed instrument)**

Name: _____

Date: _____

Quarter: _____

1. **This quarter I learned more about and focused on ……**(list 2 strategies or concepts you improved or learned about and explain) - Guskey level 2

   _____
   _____
   _____
   _____
   _____
   _____
   _____
   _____
   _____
   _____

2. **This quarter I felt/did not feel supported through…….** (reflect on any organizational support you have or have not received such as administration, IDT's, other peers) - Guskey level 3

   _____
   _____
   _____
   _____
   _____
   _____
   _____
   _____
   _____

3. **This quarter I implemented the following……**(reflect on any strategies you implemented and describe your experience) - Guskey level 4

_____
_____
_____
_____
_____
_____
_____
_____
_____

4. **This quarter my students improved in the following areas…..** (describe student learning as a result of your instructional changes) - Guskey level 5

_____
_____
_____
_____
_____
_____
_____
_____
_____

## Appendix F

**Instrument # 5: SHS Math classroom visitation protocol (Researcher developed protocol)**

Date: _____ Class: _____

Classroom:

_____

Visitor:

_____

**Students**

____ Mathematical Practice 1:Make sense of problems and persevere in solving them.
____ Mathematical Practice 2: Reason abstractly and quantitatively.
____ Mathematical Practice 3: Construct viable arguments and critique the reasoning of others.
____ Mathematical Practice 4: Model with mathematics.
____ Mathematical Practice 5: Use appropriate tools strategically.
____ Mathematical Practice 6: Attend to precision.
____ Mathematical Practice 7: Look for and make use of structure.
____ Mathematical Practice 8: Look for and express regularity in repeated reasoning.

**Teachers**

□ Promote perseverance
□ Promote student to student discourse
□ Assign purposeful group work with defined roles
□ Ask prompting, open ended questions
□ Give student choice

**Additional observations:**

_____
_____
_____
_____
_____
_____
_____

## Appendix G

**Instrument #6: Math classroom observation guide**

| | PREPARING | GETTING STARTED | MOVING ALONG | IN PLACE |
|---|---|---|---|---|
| 1. Instructional Shifts | | | | |
| **Focus/ Coherence:** Alignment of Content | ❑ None of the content in the lesson is found in the appropriate grade level standards.<br>❑ Learning intentions/targets and success criteria are not posted. | ❑ Some of the content in the lesson is found in the appropriate grade level standards.<br>❑ Learning intentions/targets and success criteria are posted but not tied to the CCSS. | ❑ Most of the content in the lesson is found in the appropriate grade level standards.<br>❑ Learning intentions/targets and success criteria are posted and tied to the CCSS. | ❑ All of the content in the lesson is found in the appropriate grade level standards.<br>❑ Learning intentions/targets and success criteria are posted, clearly tied to the CCSS, and used during the lesson. |
| **Focus/ Coherence:** Connections | ❑ The content of the lesson is not connected to the major mathematical topics at the grade level.<br>❑ There are no connections to other grade level content. | ❑ The content of the lesson is minimally connected to the major mathematical topics at the grade level.<br>❑ There are only tangential connections to other grade level content. | ❑ The content of the lesson is moderately connected to the major mathematical topics at the grade level.<br>❑ There are some connections to other grade level content. | ❑ The content of the lesson is clearly connected to the major mathematical topics at the grade level.<br>❑ There are strong connections to other grade level content. |
| **Rigor:** Cognitive demand of lesson content | ❑ The content of the lesson is not conceptually demanding for students.<br>❑ The lesson focuses on memorization of mathematical facts and procedures. | ❑ The content of the lesson is somewhat conceptually demanding.<br>❑ The lesson may introduce conceptual understanding but focuses primarily on practicing procedures during learning activities.<br>❑ Teacher asks low level questions and does not require students to explain their thinking. | ❑ The content of the lesson is conceptually demanding.<br>❑ The mathematics involved is primarily conceptual in nature or involves procedures with explicit underlying conceptual connections.<br>❑ Teacher asks a mix of higher and lower level questions that limit students opportunity to explain their thinking. | ❑ The content of the lesson is very conceptually demanding.<br>❑ The teacher maintains high cognitive demand throughout the lesson, requiring students to deeply engage with making sense of the mathematics and justifying their thinking.<br>❑ Teacher consistently asks higher-level questions that require students to explain their thinking.<br>❑ Teacher begins lesson with a review of critical prerequisite skills and concepts. |
| | PREPARING | GETTING STARTED | MOVING ALONG | IN PLACE |

| | | | | |
|---|---|---|---|---|
| **Rigor:** Procedural Skill and Fluency | ❑ Few students know the procedural skills needed to solve mathematical problems.<br>❑ Students demonstrate a lack of fluency of math facts. | ❑ Some students have learned procedural skills.<br>❑ Students have limited fluency of math facts and are slow when solving mathematical problems. | ❑ Many students have learned procedural skills.<br>❑ Students are fluent in their math facts but unable to use those facts effectively within higher-level procedures and/or when solving problems of a conceptual nature. | ❑ Most students have learned the procedural skills required by the Standards.<br>❑ Students demonstrate fluency of math facts and are able to apply those facts to higher-level procedures and mathematical thinking when problem-solving. |
| **Rigor:** Application | ❑ Teacher makes no connection between the topic of the lesson and real world situations. | ❑ Teacher makes some attempts to connect the topic of the lesson with real world situations. | ❑ Teachers consistently makes connections between the topic of the lesson and real world situations. | ❑ Teachers and students co-construct the application of math concepts to real world situations. |
| 2. Mathematical Practices | | | | |

| | | | | |
|---|---|---|---|---|
| Making sense of problems & persevering in solving them | ❑ Not observed | ❑ Limited and only tangential attention or use is more of an afterthought. | ❑ Some evidence of use, but inconsistent, missed opportunities to use or without focus or emphasis. | ❑ Teachers take every opportunity to develop number sense by ask for estimates, mental calculations, and equivalent forms of numbers.<br>❑ Students persevere in solving difficult and worthwhile problems.<br>❑ Teachers elicit, value, and celebrate alternative approaches to solving problems; students are taught that mathematics is a sense making process for understanding. |
| Reason abstractly and quantitatively | ❑ Not observed | ❑ Limited and only tangential attention or use is more of an afterthought. | ❑ Some evidence of use, but inconsistent, missed opportunities to use or without focus or emphasis. | ❑ Students make sense of quantities and their relationships in problem situations<br>❑ Student use varied representations and approaches when solving problems.<br>❑ Students know and flexibly use different properties of operations and objects. |
| | **PREPARING** | **GETTING STARTED** | **MOVING ALONG** | **IN PLACE** |
| Construct viable arguments and critique the reasoning of others | ❑ Not observed | ❑ Limited and only tangential attention or use is more of an afterthought. | ❑ Some evidence of use, but inconsistent, missed opportunities to use or without focus or emphasis. | ❑ Students explain their thinking.<br>❑ Students build upon their own and others' thinking.<br>❑ Students critique the arguments and reasoning of others. |
| Model with mathematics | ❑ Not observed | ❑ Limited and only tangential attention or use is more of an afterthought.<br>❑ | ❑ Some evidence of use, but inconsistent, missed opportunities to use or without focus or emphasis. | ❑ Students apply the mathematics they know to solve problems arising in everyday life and the workplace.<br>❑ Students analyze mathematical relationships to draw conclusions.<br>❑ Students can apply what they know and are comfortable making assumptions and approximations. |

| | | | | |
|---|---|---|---|---|
| Use appropriate tools strategically | ❑ Not observed | ❑ Limited and only tangential attention or use is more of an afterthought. | ❑ Some evidence of use, but inconsistent, missed opportunities to use or without focus or emphasis. | ❑ Teachers provide multiple<br>❑ Teachers provide multiple representations (models, number lines, tables, graphs, as well as symbols) to support visualization of skills and concepts.<br>❑ Students consider the available tools when solving a mathematical problem. |
| Attend to precision | ❑ Not observed | ❑ Limited and only tangential attention or use is more of an afterthought. | ❑ Some evidence of use, but inconsistent, missed opportunities to use or without focus or emphasis. | ❑ Teachers and students use mathematical terminology and vocabulary with precision. |

| | | | | |
|---|---|---|---|---|
| Look for and make sense of structure | ❑ Not observed | ❑ Limited and only tangential attention or use is more of an afterthought. | ❑ Some evidence of use, but inconsistent, missed opportunities to use or without focus or emphasis. . | ❑ Students can look for, develop, generalize and describe a pattern orally, symbolically, graphically, and in written form. |
| Look for regularity in repeated reasoning | ❑ Not observed | ❑ Limited and only tangential attention or use is more of an afterthought. ❑ | ❑ Some evidence of use, but inconsistent, missed opportunities to use or without focus or emphasis. | ❑ Students notice if calculations are repeated and look both for general methods and for short cuts. ❑ Students maintain oversight of the mathematical process while attending to detail. |
| | **PREPARING** | **GETTING STARTED** | **MOVING ALONG** | **IN PLACE** |
| 3. Level of Student Engagement | | | | |
| | ❑ Teacher does not appear to have control of classroom management. ❑ Few students are on task during the course of the lesson. ❑ Lesson allows students little opportunity to engage with the lesson content. ❑ Teacher does not facilitate any classroom discussion among students. | ❑ Class is organized and routines are evident. ❑ Some students are on task; others are off-task and some are being disruptive. ❑ Lesson allows students some opportunity to engage with the lesson content. ❑ Students have some opportunities to participate in classroom discussions. | ❑ Class routines are clearly established and followed. ❑ Most students are actively engaged in lesson activities. Some students may be off task when working independently. ❑ Most aspects of the lesson provide students with opportunities to engage with the lesson's central content. ❑ Students make some connections to others' thinking during classroom discussions. | ❑ Class routines are working effectively to facilitate learning. ❑ All students are actively engaged in lesson activities. Students are on task even when working independently. ❑ All aspects of the lesson provide opportunities for students to engage with the central content. ❑ Students engage in productive classroom discussions making connections and collaborating with others. ❑ Teachers have carefully planned tasks, activities, questions, and assessments for coherence. |
| 4. Differentiation | | | | |

| | | | | |
|---|---|---|---|---|
| | ❑ There is little evidence of differentiation. Instruction is oriented toward a single level of difficulty.<br>❑ There is little evidence of varied learning activities/instructional strategies.<br>❑ There is little evidence of varied assessment methods. | ❑ Instruction is differentiated for some students. Instruction is primarily oriented toward a single level of difficulty.<br>❑ Different learning activities/instructional strategies are provided for groups.<br>❑ Some different assessment methods are used. | ❑ Instruction is differentiated for all students. Instruction is oriented to more than one level of difficulty.<br>❑ Some choices are available for student learning. A range of different instructional strategies is used.<br>❑ Different assessment methods are offered to students. | ❑ Instruction is oriented toward multiple levels of difficulty. All students are working at an appropriate level.<br>❑ Multiple activities and choices are available for student learning. A comprehensive set of strategies is used.<br>❑ Students choose how they will demonstrate their learning.<br>❑ Teachers use formative assessment to consistently to check on lesson objectives/targets. |
| | **PREPARING** | **GETTING STARTED** | **MOVING ALONG** | **IN PLACE** |
| 5. Classroom Environment | | | | |

| | | | | |
|---|---|---|---|---|
| | ❑ Room contains mostly generic educational posters. No actual student work is posted in the room.<br>❑ Room contains limited resources (e.g. word walls, academic language, procedural explanations) for students. | ❑ Examples of student work are posted in the classroom, but many are outdated or with no teacher commentary or connection to the standards.<br>❑ Room contains some resources (e.g. word walls, academic language, procedural explanations) that can be used by students | ❑ Reasonably current student work is posted in the classroom with some teacher commentary. No connections to the Standards are evident.<br>❑ Room contains multiple resources that can be used by students. | ❑ Current student work is posted in the classroom with teacher commentary. Teacher comments show connections to the Standards.<br>❑ Room contains multiple resources that can be used by students AND there is evidence that students regularly access these resources. |
| **6. Culturally Responsive Teaching** | | | | |
| Students' lives | ❑ No evidence of students' lives, interests, families, communities and/or cultures are connected to the standards being taught. | ❑ Little evidence of students' lives, interests, families, communities and/or cultures are connected to the standards being taught. | ❑ Some evidence of students' lives, interests, families, communities and/or cultures are connected to the standards being taught. | ❑ Strong evidence that students' lives, interests, families, communities and/or cultures are connected to the standards being taught. |
| Diverse experiences | ❑ Delivery of content does not support diverse experiences and perspectives. | ❑ Delivery of content occasionally supports diverse experiences and perspectives. | ❑ Delivery of content inconsistently supports diverse experiences and perspectives. | ❑ Delivery of content is consistently supports diverse experiences and perspectives. |
| Respect and rapport | ❑ Limited evidence of respect and rapport among students and between teacher and students. | ❑ Some evidence of respect and rapport is among students and between teacher and students. | ❑ Most interactions among students and between teacher and students are positive and supportive. | ❑ Interactions among students and between teacher and students are consistently positive and supportive. |

## Appendix H

**Instrument #7: Summative student survey math**
**(Researcher developed instrument)**

Today's date:
Course:
Grade:

**Transfer**
**(Guskey level 4)**
**How frequently do you experience the following scenarios?**
(from Instructional Implementation Sequence: Attaining the CCSS Mathematical Practices Engagement Strategies)
(Every lesson - 4, Regularly - 3, Sometimes - 2, Never - 1)

We think and share in pairs
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

I can show my thinking in classrooms
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

I am given time to answer when questioned
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

I am allowed to struggle and encouraged to persevere
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

I am encouraged to explain my reasoning
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

I am consistently asked higher- level questions that require me to explain my thinking.
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

Lessons begin with a review of critical prerequisite skills and concepts.
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

My teacher and I work together to apply math
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

I am asked for estimates, mental calculations, and equivalent forms of numbers.
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

I am encouraged to use alternative approaches to solving problems; I am taught that mathematics is a sense making process for understanding.
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

My teacher provides multiple representations (models, number lines, tables, graphs, as well as symbols) to support visualization of skills and concepts.
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

Our class routines are working effectively to facilitate learning.
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

I am actively engaged in lesson activities. I am on task even when working independently.
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes

☐ Never

All aspects of our lessons provide opportunities for me to engage with the content.
☐ Every lesson
☐ Regularly
☐ Sometimes
☐ Never

I engage in productive classroom discussions making connections and collaborating with others.
☐ Every lesson
☐ Regularly
☐ Sometimes
☐ Never

The work that is given to me in Algebra class is at an appropriate level for me.
☐ Every lesson
☐ Regularly
☐ Sometimes
☐ Never

Multiple activities and choices are available for my learning. A comprehensive set of strategies is used.
☐ Every lesson
☐ Regularly
☐ Sometimes
☐ Never

I am allowed to choose how I will demonstrate my learning.
☐ Every lesson
☐ Regularly
☐ Sometimes
☐ Never

**Student learning**
**(Guskey level 5)**

**I am doing this on a regular basis during class:**
(every lesson/regularly/sometimes/never)
(from GE CCSS Math classroom observation guide)
(Every lesson - 4, Regularly - 3, Sometimes - 2, Never - 1)

**1. Make sense of problems and persevere in solving them**
I persevere in solving difficult and worthwhile problems
☐ Every lesson
☐ Regularly
☐ Sometimes
☐ Never

**2. Reason abstractly and quantitatively.**

I make sense of quantities and their relationships in problem situations

- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

I use varied representations and approaches when solving problems
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

I know and flexibly use different properties of operations and objects
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

**3. Construct viable arguments and critique the reasoning of others.**

I explain my thinking
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

I build upon my own and others' thinking
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

I critique the arguments and reasoning of others
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

**4. Model with mathematics.**

I apply the mathematics I know to solve problems arising in everyday life and the workplace
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

I analyze mathematical relationships to draw conclusions
- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

I apply what I know and am comfortable making assumptions and approximations

- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

**5. I use appropriate tools strategically**

I consider the available tools when solving a mathematical problem

- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

**6. I attend to precision.**

I use mathematical terminology and vocabulary with precision

- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

**7. I look for and make use of structure.**

I look for, develop, generalize, and describe a pattern orally, symbolically, graphically, and in written form

- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

**8. Look for and make use of regularity in repeated reasoning.**

I look for and make use of regularity in repeated reasoning.

- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

I maintain oversight of the mathematical process while attending to detail.

- ☐ Every lesson
- ☐ Regularly
- ☐ Sometimes
- ☐ Never

**33 questions**

## Appendix I
## Exploratory Factor Analysis Results

*Survey Questions*

| # | Question | MP | Level |
|---|---|---|---|
| 1 | We think and share in pairs in Algebra class. | MP 1, 3 | 4 |
| 2 | I can show my thinking. | MP 3, 6 | 4 |
| 3 | I am given time to answer when questioned. | MP 1, 3 | 4 |
| 4 | I am allowed to struggle and encouraged to persevere. | MP 1 | 4 |
| 5 | I am encouraged to explain my reasoning. | MP 2, 8 | 4 |
| 6 | I am consistently asked higher-level questions that require me to explain my thinking. | IS | 4 |
| 7 | Lessons begin with a review of critical prerequisite skills and concepts. | IS | 4 |
| 8 | My teacher and I work together to apply math concepts to real world situations. | IS | 4 |
| 9 | I am asked for estimates, mental calculations, and equivalent forms of numbers. | MP 1 | 4 |
| 10 | I am encouraged to use alternative approaches to solving problems; I am taught that mathematics is a sense making process for understanding. | MP 1 | 4 |
| 11 | My teacher provides multiple representations (models, number lines, tables, graphs, as well as symbols) to support understanding of skills and concepts. | MP 5 | 4 |
| 12 | Our class routines are working effectively to facilitate learning. | SE | 4 |
| 13 | I am actively engaged in lesson activities. I am on task even when working independently. | SE | 4 |
| 14 | All aspects of our lessons provide opportunities for me to engage with the content. | SE | 4 |
| 15 | I engage in productive classroom discussions making connections and collaborating with others. | SE | 4 |
| 16 | The work that is given to me in Algebra class is at an appropriate level for me. | SE | 4 |
| 17 | Multiple activities and choices are available for my learning during Algebra class. A comprehensive set of strategies is used. | SE | 4 |
| 18 | I am allowed to choose how I will demonstrate my learning. | SE | 4 |
| 19 | I persevere in solving difficult and worthwhile problems. | MP 1 | 5 |
| 20 | I make sense of quantities and their relationships in problem situations. | MP 2 | 5 |
| 21 | I use varied representations and approaches when solving problems. | MP 2 | 5 |
| 22 | I know and flexibly use different properties of operations and objects. | MP 2 | 5 |
| 23 | I explain my thinking. | MP 3 | 5 |

| | | | |
|---|---|---|---|
| 24 | I build upon my own and others' thinking. | MP 3 | 5 |
| 25 | I critique the arguments and reasoning of others. | MP 3 | 5 |
| 26 | I apply the mathematics I know to solve problems arising in everyday life and the workplace. | MP 4 | 5 |
| 27 | I analyze mathematical relationships to draw conclusions. | MP 4 | 5 |
| 28 | I apply what I know and am comfortable making assumptions and approximations. | MP 4 | 5 |
| 29 | I consider the available tools when solving a mathematical problem. | MP 5 | 5 |
| 30 | I use mathematical terminology and vocabulary with precision. | MP 6 | 5 |
| 31 | I look for, develop, generalize, and describe a pattern orally, symbolically, graphically, and in written form. | MP 7 | 5 |
| 32 | I look for and make use of regularity in repeated reasoning. | MP 8 | 5 |
| 33 | I maintain oversight of the mathematical process while attending to detail. | MP 8 | 5 |

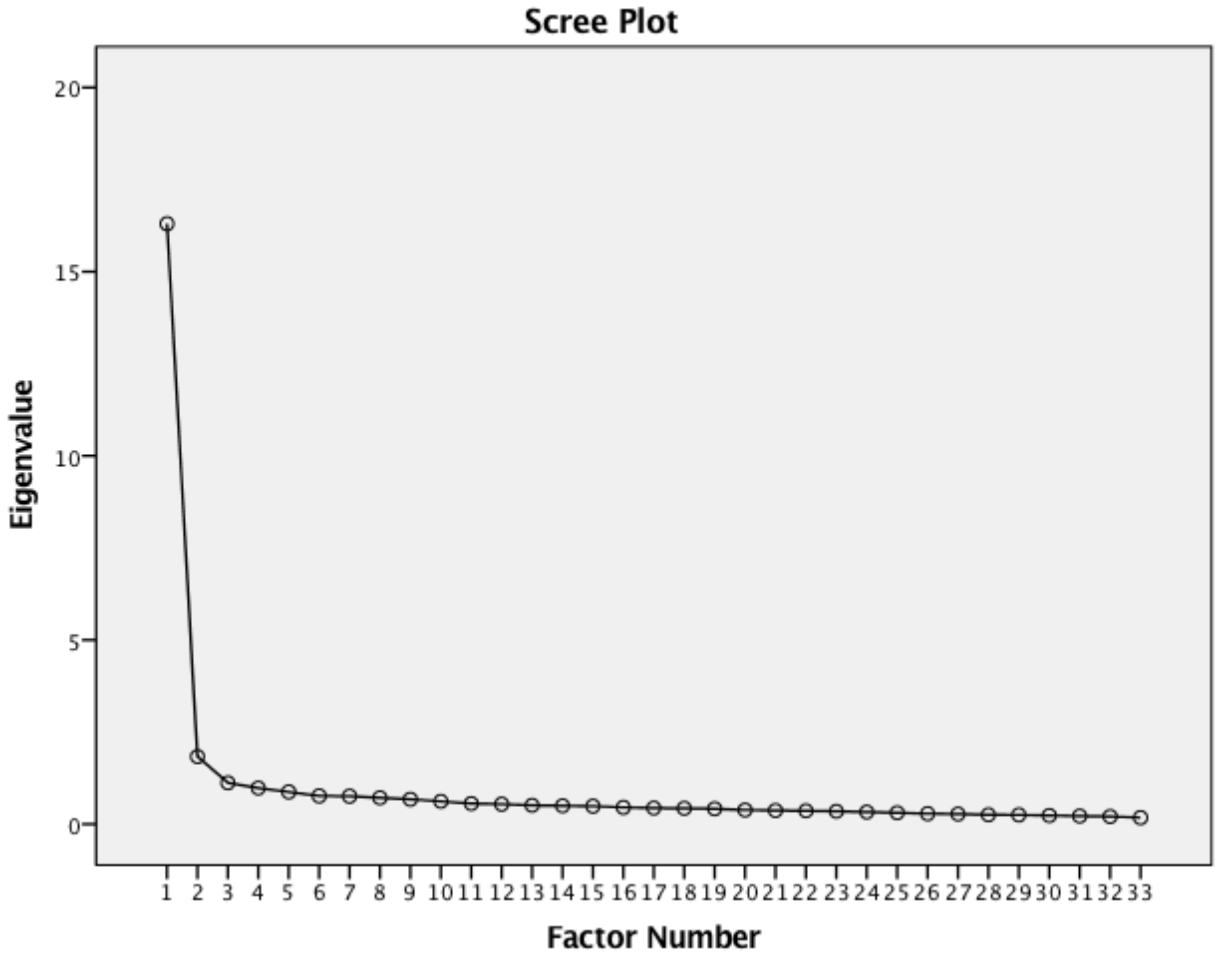
*Total Variance Explained*

| | Initial Eigenvalues | | | Extraction Sums of Squared Loadings | | | Rotation Sums of Squared Loadings[a] |
|---|---|---|---|---|---|---|---|
| Factor | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % | Total |
| 1 | 16.302 | 49.401 | 49.401 | 15.847 | 48.022 | 48.022 | 14.689 |
| 2 | 1.835 | 5.561 | 54.962 | 1.392 | 4.217 | 52.239 | 13.900 |
| 3 | 1.123 | 3.404 | 58.366 | .609 | 1.847 | 54.086 | .768 |
| 4 | .980 | 2.970 | 61.336 | | | | |
| 5 | .874 | 2.648 | 63.983 | | | | |
| 6 | .766 | 2.321 | 66.304 | | | | |
| 7 | .760 | 2.303 | 68.607 | | | | |
| 8 | .714 | 2.163 | 70.770 | | | | |
| 9 | .676 | 2.049 | 72.819 | | | | |
| 10 | .618 | 1.874 | 74.693 | | | | |
| 11 | .561 | 1.700 | 76.393 | | | | |
| 12 | .542 | 1.643 | 78.035 | | | | |
| 13 | .511 | 1.549 | 79.584 | | | | |
| 14 | .499 | 1.511 | 81.095 | | | | |
| 15 | .488 | 1.478 | 82.573 | | | | |
| 16 | .454 | 1.377 | 83.950 | | | | |
| 17 | .439 | 1.329 | 85.279 | | | | |
| 18 | .429 | 1.300 | 86.579 | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 19 | .419 | 1.269 | 87.848 | | | | |
| 20 | .385 | 1.168 | 89.016 | | | | |
| 21 | .370 | 1.122 | 90.138 | | | | |
| 22 | .361 | 1.095 | 91.233 | | | | |
| 23 | .348 | 1.055 | 92.287 | | | | |
| 24 | .330 | 1.001 | 93.289 | | | | |
| 25 | .311 | .941 | 94.230 | | | | |
| 26 | .286 | .865 | 95.095 | | | | |
| 27 | .275 | .833 | 95.929 | | | | |
| 28 | .253 | .765 | 96.694 | | | | |
| 29 | .251 | .760 | 97.454 | | | | |
| 30 | .235 | .713 | 98.167 | | | | |
| 31 | .220 | .667 | 98.834 | | | | |
| 32 | .208 | .631 | 99.465 | | | | |
| 33 | .177 | .535 | 100.000 | | | | |

Extraction Method: Maximum Likelihood.

*Scree Plot*



*Pattern Matrix*

| | Factor | | |
|---|---|---|---|
| | 1 | 2 | 3 |
| question27 | .805 | | |
| question30 | .796 | | |
| question19 | .791 | | |
| question32 | .780 | | |
| question28 | .773 | | |
| question31 | .769 | | |
| question21 | .744 | | |
| question20 | .739 | | |
| question29 | .734 | | -.178 |
| question22 | .722 | | |
| question24 | .721 | | .204 |
| question25 | .707 | | .188 |
| question33 | .678 | | -.156 |
| question26 | .678 | | |
| question23 | .676 | | .138 |
| question13 | .389 | .277 | -.130 |
| question11 | -.129 | .859 | |
| question10 | | .777 | |
| question8 | | .751 | .127 |
| question12 | | .719 | -.207 |
| question9 | | .679 | .227 |
| question7 | | .677 | |
| question18 | | .667 | -.152 |
| question4 | | .630 | -.119 |
| question1 | | .624 | .106 |
| question3 | | .609 | -.105 |
| question17 | .166 | .581 | -.210 |
| question5 | | .562 | |
| question15 | .225 | .511 | .131 |
| question14 | .312 | .468 | -.130 |
| question16 | .179 | .461 | -.336 |
| question6 | .222 | .370 | .240 |
| question2 | .314 | .355 | |

## Appendix J
## Instructional Coach Survey for Faculty Results

*Instructional Coach Faculty Survey Scores*

| Item | Mean Score |
|---|---|
| This instructional coach meets with me during the scheduled time. | 3.58 |
| This instructional coach helps me overcome barriers to teaching and learning. | 2.75 |
| This instructional coach helps me evaluate my instruction. | 3.08 |
| This instructional coach helps me to identify and solve problems. | 2.83 |
| This instructional coach finds new ways to do things better. | 2.42 |
| This instructional coach contributes positively to improvement of instruction. | 3.00 |
| This instructional coach believes in celebrating instructional and academic improvement. | 2.73 |
| This instructional coach maintains open, two-way communication with school faculty. | 2.75 |
| This instructional coach sets high standards for teacher performance. | 3.55 |
| This instructional coach performs duties in a professional manner | 3.50 |
| This instructional coach communicates information clearly and succinctly. | 3.42 |
| This instructional coach communicates the importance of focusing on the needs of students. | 3.08 |
| This instructional coach promotes and supports innovations. | 3.25 |
| This instructional coach is a positive person who maintains his/her composure. | 3.25 |
| This instructional coach helps me establish routines and procedures that contribute to learning and teaching of Math content. | 2.08 |
| This instructional coach conducts planning and feedback sessions with me. | 3.00 |
| This instructional coach works with me to ensure that standards and instruction are aligned. | 1.67 |
| This instructional coach provides feedback to me about my instruction. | 3.58 |
| This instructional coach assists me in instructional planning when needed. | 2.10 |
| This instructional coach understands research-based instructional procedures and helps me implement these procedures. | 2.70 |
| This instructional specialist keeps abreast of teaching/learning research and best practices. | 2.70 |
| This instructional coach is a mentor to me. | 2.25 |
| | |
| Total average score | 2.88 |
| Algebra 1 average score | 3.33 |
| Algebra 2 average score | 2.68 |

*Mean Score Per Teacher*

| Teacher | Average Score |
|---|---|
| 1 | 2.77 |
| 2 | 3.57 |
| 3 | 1.32 |
| 4 | 3.00 |
| 5 | 2.23 |
| 6 | 2.00 |
| 7 | 3.64 |
| 8 | 3.50 |
| 9 | 4.00 |
| 10 | 3.45 |
| 11 | 2.24 |
| 12 | 3.00 |

## Appendix K
## Teacher Survey Pre and Post Results

Mean Scores of Teachers' Learning

| **Survey items** | $\overline{x}$ **Pre** | $\overline{x}$**Post** | **Diff.** |
|---|---|---|---|
| The instructional shifts in mathematics. | 2.75 | 3.08 | 0.33 |
| The maintenance of a high cognitive demand throughout the lessons, requiring students to deeply engage with making sense of the mathematics and justifying their thinking. | 2.75 | 3.17 | 0.42 |
| The asking of higher- level questions that require students to explain their thinking. | 3.00 | 3.33 | 0.33 |
| The review of critical prerequisite skills and concepts. | 3.33 | 3.58 | 0.25 |
| The application of math concepts to real world situations. | 3.25 | 3.25 | 0.00 |
| The eight Mathematical Practices | 2.75 | 3.00 | 0.25 |
| The development of number sense by asking for estimates, mental calculations, and equivalent forms of numbers. | 3.17 | 3.67 | 0.50 |
| The value and promotion of alternative approaches to solving problems; students are taught that mathematics is a sense making process for understanding. | 3.08 | 3.42 | 0.33 |
| The provision of multiple representations (models, number lines, tables, graphs, as well as symbols) to support visualization of skills and concepts. | 3.33 | 3.83 | 0.50 |
| The use of mathematical terminology and vocabulary with precision. | 3.42 | 3.58 | 0.17 |
| Class routines that are working effectively to facilitate learning. | 3.25 | 3.50 | 0.25 |
| Students' active engagement in lesson activities. Students are on task even when working independently. | 3.00 | 3.42 | 0.42 |
| The provision of opportunities for students to engage with the central content within all aspects of the lesson | 2.50 | 2.92 | 0.42 |
| Students' engagement in productive classroom discussions making connections and collaborating with others. | 2.92 | 3.17 | 0.25 |
| The careful planning of tasks, activities, questions, and assessments for coherence. | 3.08 | 3.17 | 0.08 |
| Instruction's orientation toward multiple levels of difficulty. All students are working at an appropriate level. | 2.58 | 2.83 | 0.25 |
| The availability of multiple activities and choices for student learning. A comprehensive set of strategies is used. | 2.50 | 3.00 | 0.50 |
| Students' choice how they will demonstrate their learning. | 2.25 | 2.75 | 0.50 |
| The use of formative assessment to consistently check on lesson objectives/targets. | 2.92 | 3.08 | 0.17 |
| **Level 2 total average** | **2.94** | **3.25** | **0.31** |

*Note:* Teacher survey, Guskey (2000) Level 2, teacher learning, mean scores of pre and post test.

Mean Scores of Teachers' Transfer of Knowledge and Skills.

| **Survey items** | $\bar{x}$ **Pre** | $\bar{x}$ **Post** | **Diff.** |
|---|---|---|---|
| I initiate think, pair share | 2.33 | 2.42 | 0.08 |
| I require students' to show (communicate) their thinking in classrooms | 3.33 | 3.58 | 0.25 |
| I question and wait | 3.42 | 3.50 | 0.08 |
| I use grouping and engaging problems | 2.67 | 2.75 | 0.08 |
| I use questions and prompts with groups | 2.75 | 2.92 | 0.17 |
| I allow students to struggle and encourage them to persevere | 3.17 | 3.50 | 0.33 |
| I encourage students to explain their reasoning | 3.25 | 3.75 | 0.50 |
| I maintain high cognitive demand throughout my lessons, requiring students to deeply engage with making sense of the mathematics and justifying their thinking. | 2.92 | 3.42 | 0.50 |
| I consistently ask higher- level questions that require students to explain their thinking. | 2.92 | 3.25 | 0.33 |
| I begin my lessons with a review of critical prerequisite skills and concepts. | 2.58 | 2.92 | 0.33 |
| My students and I co-construct the application of math concepts to real world situations. | 1.92 | 2.25 | 0.33 |
| I take every opportunity to develop number sense by ask for estimates, mental calculations, and equivalent forms of numbers. | 2.58 | 2.83 | 0.25 |
| I elicit, value, and celebrate alternative approaches to solving problems; students are taught that mathematics is a sense making process for understanding. | 3.00 | 3.33 | 0.33 |
| I provide multiple representations (models, number lines, tables, graphs, as well as symbols) to support visualization of skills and concepts. | 3.17 | 3.33 | 0.17 |
| I use mathematical terminology and vocabulary with precision. | 3.42 | 3.58 | 0.17 |
| My class routines are working effectively to facilitate learning. | 3.33 | 3.25 | -0.08 |
| All my students are actively engaged in lesson activities. Students are on task even when working independently. | 2.83 | 3.00 | 0.17 |
| All aspects of my lesson provide opportunities for students to engage with the central content. | 2.83 | 2.92 | 0.08 |
| My students engage in productive classroom discussions making connections and collaborating with others. | 2.83 | 2.92 | 0.08 |
| I have carefully planned tasks, activities, questions, and assessments for coherence. | 3.00 | 3.33 | 0.33 |
| My instruction is oriented toward multiple levels of difficulty. All students are working at an appropriate level. | 2.50 | 3.00 | 0.50 |
| I make multiple activities and choices available for student learning. A comprehensive set of strategies is used. | 2.25 | 2.58 | 0.33 |
| I allow students to choose how they will demonstrate their learning. | 2.42 | 2.17 | -0.25 |
| I use formative assessment consistently to check on lesson objectives/targets. | 3.08 | 2.92 | -0.17 |
| **Level 4 total average** | **2.85** | **3.06** | **0.21** |

*Note*: Teacher survey pre-post Guskey (2000) level 4, mean scores of transfer of knowledge and skills.

Mean Scores for Student Learning Outcomes

| **Survey items** | $\bar{x}$ **Pre** | $\bar{x}$ **Post** | **Diff.** |
|---|---|---|---|
| Persevere in solving difficult and worthwhile problems | 2.67 | 2.92 | 0.25 |
| Make sense of quantities and their relationships in problem situations | 2.67 | 2.92 | 0.25 |
| Use varied representations and approaches when solving problems | 2.50 | 2.92 | 0.42 |
| Know and flexibly use different properties of operations and objects | 2.67 | 2.92 | 0.25 |
| Explain their thinking | 2.58 | 2.92 | 0.33 |
| Build upon their own and others' thinking | 2.83 | 3.00 | 0.17 |
| Critique the arguments and reasoning of others | 2.50 | 2.75 | 0.25 |
| Apply the mathematics they know to solve problems arising in everyday life and the workplace | 2.50 | 2.58 | 0.08 |
| Analyze mathematical relationships to draw conclusions | 2.50 | 2.50 | 0.00 |
| Apply what they know and are comfortable making assumptions and approximations | 2.67 | 2.58 | -0.08 |
| Consider the available tools when solving a mathematical problem | 3.00 | 3.17 | 0.17 |
| Use mathematical terminology and vocabulary with precision | 2.58 | 2.92 | 0.33 |
| Look for, develop, generalize, and describe a pattern orally, symbolically, graphically, and in written form | 2.50 | 2.67 | 0.17 |
| Notice if calculations are repeated and look both for general methods and for short cuts | 2.42 | 2.67 | 0.25 |
| Maintain oversight of the mathematical process while attending to detail. | 2.25 | 2.67 | 0.42 |
| **Level 5 total average** | **2.59** | **2.81** | **0.22** |

*Note*: Teacher survey pre-post Guskey (2000) level 5, mean scores for student learning outcomes

## Appendix L
## Visitation Protocol Results

Visitation Protocol Frequencies

| **Observed teaching practice** | **Frequency of observed use** | **% use** |
|---|---|---|
| Promote independence and perseverance | 10 | 30% |
| Promote student to student discourse | 11 | 33% |
| Assign purposeful group work with defined roles | 9 | 27% |
| Ask prompting, open ended questions on all levels of Webb's Depth of Knowledge esp. 3 and 4 | 14 | 42% |
| Vary ways in which a student can achieve a standard | 9 | 27% |
| Model | 17 | 52% |

Visitation Protocol Frequencies

| **Observed Item** | **Frequency** | **%** |
|---|---|---|
| Mathematical Practice 1:Make sense of problems and persevere in solving them. | 12 | 36% |
| Mathematical Practice 2: Reason abstractly and quantitatively. | 8 | 24% |
| Mathematical Practice 3: Construct viable arguments and critique the reasoning of others. | 6 | 18% |
| Mathematical Practice 4: Model with mathematics. | 15 | 45% |
| Mathematical Practice 5: Use appropriate tools strategically. | 10 | 30% |
| Mathematical Practice 6: Attend to precision. | 16 | 48% |
| Mathematical Practice 7: Look for and make use of structure. | 8 | 24% |
| Mathematical Practice 8: Look for and express regularity in repeated reasoning. | 5 | 15% |

## Appendix M
## Student Survey Results

Mean Scores of Experienced Teaching Strategies

| **Survey items** | $\overline{x}$ |
|---|---|
| We think and share in pairs in Algebra class. | 2.48 |
| I can show my thinking. | 2.75 |
| I am given time to answer when questioned. | 2.85 |
| I am allowed to struggle and encouraged to persevere. | 2.80 |
| I am encouraged to explain my reasoning. | 2.91 |
| I am consistently asked higher-level questions that require me to explain my thinking. | 2.72 |
| Lessons begin with a review of critical prerequisite skills and concepts. | 2.60 |
| My teacher and I work together to apply math concepts to real world situations. | 2.48 |
| I am asked for estimates, mental calculations, and equivalent forms of numbers. | 2.72 |
| I am encouraged to use alternative approaches to solving problems; I am taught that mathematics is a sense making process for understanding | 2.74 |
| My teacher provides multiple representations (models, number lines, tables, graphs, as well as symbols) to support understanding of skills and concepts. | 2.99 |
| Our class routines are working effectively to facilitate learning. | 2.74 |
| I am actively engaged in lesson activities. I am on task even when working independently. | 2.87 |
| All aspects of our lessons provide opportunities for me to engage with the content. | 2.79 |
| I engage in productive classroom discussions making connections and collaborating with others. | 2.70 |
| The work that is given to me in Algebra class is at an appropriate level for me. | 2.95 |
| Multiple activities and choices are available for my learning during Algebra class. A comprehensive set of strategies is used. | 2.74 |
| I am allowed to choose how I will demonstrate my learning. | 2.62 |
| **Level 4 total average** | **2.75** |

*Note:* Student survey Guskey (2000) level 4, mean scores of experienced teaching strategies

Mean Scores of Experienced Teaching Practices

| **Survey item from teacher's perspective** | $\overline{x}$ **T post** | $\overline{x}$ **Stud.** | **Survey item from student's perspective** |
|---|---|---|---|
| I initiate think, pair share | 2.42 | 2.48 | We think and share in pairs in Algebra class. |
| I require students' to show (communicate) their thinking in classrooms | 3.58 | 2.75 | I can show my thinking. |
| I question and wait | 3.50 | 2.85 | I am given time to answer when questioned. |
| I allow students to struggle and encourage them to persevere | 3.50 | 2.80 | I am allowed to struggle and encouraged to persevere. |
| I encourage students to explain their reasoning | 3.75 | 2.91 | I am encouraged to explain my reasoning. |
| I consistently ask higher- level questions that require students to explain their thinking. | 3.25 | 2.72 | I am consistently asked higher-level questions that require me to explain my thinking. |
| I begin my lessons with a review of critical prerequisite skills and concepts. | 2.92 | 2.60 | Lessons begin with a review of critical prerequisite skills and concepts. |
| My students and I co-construct the application of math concepts to real world situations. | 2.25 | 2.48 | My teacher and I work together to apply math concepts to real world situations. |
| I take every opportunity to develop number sense by ask for estimates, mental calculations, and equivalent forms of numbers. | 2.83 | 2.72 | I am asked for estimates, mental calculations, and equivalent forms of numbers. |
| I elicit, value, and celebrate alternative approaches to solving problems; students are taught that mathematics is a sense making process for understanding. | 3.33 | 2.74 | I am encouraged to use alternative approaches to solving problems; I am taught that mathematics is a sense making process for understanding |
| I provide multiple representations (models, number lines, tables, graphs, as well as symbols) to support visualization of skills and concepts. | 3.33 | 2.99 | My teacher provides multiple representations (models, number lines, tables, graphs, as well as symbols) to support understanding of skills and concepts. |
| My class routines are working effectively to facilitate learning. | 3.25 | 2.74 | Our class routines are working effectively to facilitate learning. |
| All my students are actively engaged in lesson activities. Students are on task even when working independently. | 3.00 | 2.87 | I am actively engaged in lesson activities. I am on task even when working independently. |
| All aspects of my lesson provide opportunities for students to engage with the central content. | 2.92 | 2.79 | All aspects of our lessons provide opportunities for me to engage with the content. |
| My students engage in productive classroom discussions making connections and collaborating with others. | 2.92 | 2.70 | I engage in productive classroom discussions making connections and collaborating with others. |
| My instruction is oriented toward multiple levels of difficulty. All students are working at an appropriate level. | 2.42 | 2.95 | The work that is given to me in Algebra class is at an appropriate level for me. |
| I make multiple activities and choices available for student learning. A comprehensive set of strategies is used. | 3.58 | 2.74 | Multiple activities and choices are available for my learning during Algebra class. A comprehensive set of strategies is used. |
| I allow students to choose how they will demonstrate their learning. | 2.17 | 2.62 | I am allowed to choose how I will demonstrate my learning. |
| **Level 4 total average** | **3.05** | **2.75** | |

Comparison teacher post survey and student survey Guskey (2000) level 4, mean scores of experienced teaching practices

Mean Scores Student Learning Outcomes

| **Survey items** | $\overline{x}$ |
|---|---|
| I persevere in solving difficult and worthwhile problems. | 2.79 |
| I make sense of quantities and their relationships in problem situations. | 2.75 |
| I use varied representations and approaches when solving problems. | 2.74 |
| I know and flexibly use different properties of operations and objects. | 2.77 |
| I explain my thinking. | 2.79 |
| I build upon my own and others' thinking. | 2.80 |
| I critique the arguments and reasoning of others. | 2.57 |
| I apply the mathematics I know to solve problems arising in everyday life and the workplace. | 2.55 |
| I analyze mathematical relationships to draw conclusions. | 2.68 |
| I apply what I know and am comfortable making assumptions and approximations. | 2.78 |
| I consider the available tools when solving a mathematical problem. | 2.89 |
| I use mathematical terminology and vocabulary with precision. | 2.62 |
| I look for, develop, generalize, and describe a pattern orally, symbolically, graphically, and in written form. | 2.64 |
| I look for and make use of regularity in repeated reasoning. | 2.69 |
| I maintain oversight of the mathematical process while attending to detail. | 2.72 |
| **Level 5 total average** | **2.71** |

*Note:* Student survey Guskey (2000) level 5, mean scores student learning outcomes

Mean Scores Student Learning Outcomes

| **Survey item from teacher's perspective** | $\overline{x}$ **T post** | $\overline{x}$ **Stud.** | **Survey item from student's perspective** |
|---|---|---|---|
| Persevere in solving difficult and worthwhile problems | 2.92 | 2.79 | I persevere in solving difficult and worthwhile problems. |
| Make sense of quantities and their relationships in problem situations | 2.92 | 2.75 | I make sense of quantities and their relationships in problem situations. |
| Use varied representations and approaches when solving problems | 2.92 | 2.74 | I use varied representations and approaches when solving problems. |
| Know and flexibly use different properties of operations and objects | 2.92 | 2.77 | I know and flexibly use different properties of operations and objects. |
| Explain their thinking | 2.92 | 2.79 | I explain my thinking. |
| Build upon their own and others' thinking | 3.00 | 2.80 | I build upon my own and others' thinking. |

| | | | |
|---|---|---|---|
| Critique the arguments and reasoning of others | 2.75 | 2.57 | I critique the arguments and reasoning of others. |
| Apply the mathematics they know to solve problems arising in everyday life and the workplace | 2.58 | 2.55 | I apply the mathematics I know to solve problems arising in everyday life and the workplace. |
| Analyze mathematical relationships to draw conclusions | 2.50 | 2.68 | I analyze mathematical relationships to draw conclusions. |
| Apply what they know and are comfortable making assumptions and approximations | 2.58 | 2.78 | I apply what I know and am comfortable making assumptions and approximations. |
| Consider the available tools when solving a mathematical problem | 3.17 | 2.89 | I consider the available tools when solving a mathematical problem. |
| Use mathematical terminology and vocabulary with precision | 2.92 | 2.62 | I use mathematical terminology and vocabulary with precision. |
| Look for, develop, generalize, and describe a pattern orally, symbolically, graphically, and in written form | 2.67 | 2.64 | I look for, develop, generalize, and describe a pattern orally, symbolically, graphically, and in written form. |
| Notice if calculations are repeated and look both for general methods and for short cuts | 2.67 | 2.69 | I look for and make use of regularity in repeated reasoning. |
| Maintain oversight of the mathematical process while attending to detail. | 2.67 | 2.72 | I maintain oversight of the mathematical process while attending to detail. |
| **Level 5 total average** | **2.81** | **2.71** | **Level 5 total average** |

*Note:* Comparison teacher post and Student survey Guskey (2000) level 5, mean scores student learning outcomes

**References**

Addison, R. M., & Haig, C. (2006). The performance architect's essential guide to the performance technology landscape. In J. A. Pershing (Ed). *Handbook of Human Performance Technology: Principles, Practices, and Potential* (3rd ed.). San Francisco, CA: Pfeiffer.

Allen, J., Pianta, R., Gregory, A., Mikami, A., & Lun, J. (2011). An interaction-based approach to increasing secondary school instruction and student achievement. *Science*, *333*(19), 1034-1037.

APQC Education Advanced Working Group (AWG). (2011). *Best Practices in Evaluating the Impact of Instructional Coaching on Teaching and Learning*. Final Report.

Beane, J. (2002). Beyond self-interest: A democratic core curriculum. *Educational Leadership, 59*(7), 25-28.

Biancarosa, G., Bryk, A. S., & Dexter, E. R. (2010, September). Assessing the value-added effects of Literacy Collaborative professional development on student learning. *The Elementary School Journal, 111*(1), 7-34.

Bichelmeyer, B. A., & Horvitz, B. S. (2006). Comprehensive performance evaluation: Using logic models to develop a theory-based approach for evaluation of human performance technology interventions. In J. A. Pershing (Ed). *Handbook of Human Performance Technology: Principles, Practices, and Potential* (3rd ed.). San Francisco, CA: Pfeiffer.

Coladarci, Th., Cobb, C.D., Minium, E.W., & Clarke, R.B. (2011). *Fundamentals of Statistical Reasoning in Education* (3rd ed). Hoboken, NJ: John Wiley & Sons.

Collins, K.M.T. (2010). Advanced sampling designs in mixed research: Current practices and emerging trends in the social and behavioral sciences. In A.

Tashakkori & C. Teddlie (Eds.), SAGE handbook of mixed methods in social & behavioral research (2nd ed.). Thousand Oaks, CA: Sage.

Cole, P. (2004). *Professional development:* A great way to avoid change. *Centre for Strategic Education Seminar Series Paper*, No 140.

Cooksy, L. J., Gill, O., & Kelly, P.A. (2001). The program logic model as an integrative framework for a multimethod evaluation. *Evaluation and Program Planning, 24*(2001), 119-128.

Creswell, J. (2009): *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (3rd ed.). Thousand Oaks, CA: Sage.

Cornett, J., & Knight, J. (2009). Research on coaching. In J. Knight (Ed.) *Coaching: Approaches and Perspectives.* Thousand Oaks, CA: Corwin.

Danks, S. (2011). The Addie Model: Designing evaluating instructional coach effectiveness. *ASQ Primary and Secondary Education Brief, 4*(5*).*

Desimone, L. M. (2009). Improving impact studies of teacher's professional development: Toward better conceptualizations and measures. *Educational Researcher, 38*(3), 181-199.

Desimone, L. M., Porter, A., Garet, M., Yoon, K, & Birman, B. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. *Education Evaluation and Policy Analysis, 24*(2), 81-112.

Dick, W., Carey, L., & Carey, J. O. (2015). *The Systematic Design of Instruction* (8th ed.). San Francisco, CA: Pearson.

Dick, W. (1996). The Dick and Carey model: Will it survive the decade? *Educational Technology Research Development, 44*(3), 55-63.

DuFour, R, DuFour, R., Eaker, R., & Many, T. (2010). *Learning by Doing: A Handbook for Professional Learning Communities at Work* (2nd ed). Bloomington, IN: Solution Tree Press.

Eisenhart, M., (2005). Hammers and saws for the improvement of educational research. *Educational Theory, 55*(3), 245 - 261.

Elmore, R.F. (2004). *School Reform From the Inside Out: Policy, Practice, Performance*. Cambridge, MA: Harvard Education Press.

Ferond, C. (2006). The origins and evolution of Human Performance Technology. In J. A. Pershing (Ed). *Handbook of Human Performance Technology: Principles, Practices, and Potential* (3rd ed.). San Francisco, CA: Pfeiffer.

Fogarty, R., & Pete, B. (2007). *From Staff Room to Classroom: A Guide for Planning and Coaching Professional Development*. Thousand Oaks, CA: Corwin Press.

Friendship Public Charter School (2010). *Instructional Performance coach Plan SY 2010-2011*.

Fairfax County Public Schools (2014): Instructional Coaching Program. Retrieved from http://commweb.fcps.edu/programprofile/overview.cfm?programID=207

Fullan, M. (2006). Change Theory: A force for school improvement. *Centre for Strategic Education Seminar Series Paper*, No 157.

Fullan, M. (2007). Change the terms for teacher learning. *National Staff Development Council, 28*(3), 35-36.

Garet, M. S., Cronen, S., Eaton, M., Kurki, A., Ludwig, M., Jones, W., et al. (2008). *The Impact of Two Professional Development Interventions on Early Reading Instruction and Achievement* (NCEE 2008-4030). Washington, DC: National

Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Garet, M. S., Wayne, A. J., Stancavage, F., Taylor, J., Eaton, M., Walters, K., et al. (2010). *Middle School Mathematics Professional Development Impact Study: Findings after the Second Year of Implementation. Executive Summary (*NCEE 2010-4009). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Garet, M. S., Wayne, A. J., Stancavage, F., Taylor, J., Eaton, M., Walters, K., et al. (2011). *Middle School Mathematics Professional Development Impact Study: Findings after the Second Year of Implementation. Executive Summary (*NCEE 2011-4025). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal, 38*(4), 915-945.

Garet, M., Wayne, A., Stancavage, F., Taylor, J., Walters, K., Song, M., et al. (2010). *Middle School Mathematics Professional Development Impact Study: Findings After the First Year of Implementation* (NCEE 2010-4009). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

GE Foundation Developing Futures in Educaton (2013). Evaluation of the GE Foundation-supported coaching & demonstration schools initiative in Erie Public Schools, SY 2013-2013. Working Paper.

Gosuch, R. L. (1990). Common factor analysis versus component analysis: Some well and little known facts. *Multivariate Behavioral Research, 25,* 33-39.

Greene, J. C., (2007). *Mixed Methods in Social Inquiry*. San Francisco: Josey-Bass.

Greene, J.C., Caracelli, V. J., & Graham, W.F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis, 11*(3), 255-274.

Gulamhussein, A. (2013). *Teaching the Teachers: Effective Professional Development in an Era of High Stakes Accountability.* Washington, D.C.: Center for Public Education.

Guskey, T. R., & Yoon, K. S. (2009). What works in Professional development. *Phi Delta Kappa, 90*(8), 495-500.

Guskey, T. R. (2000). *Evaluating Professional Development*. Thousand Oaks, CA: Corwin Press.

Guskey, T. R. (2002). Professional development and teacher change. *Teachers and Teaching: Theory and Practice, 8*(3/4), 381 – 391.

Guskey, T. R., & Sparks, D. (1996). Exploring the relationship between staff development and improvements in student learning. *Journal of Staff Development, 17*(4), 34–38.

Gustafson, K. L., & Branch, R. M. (2002). *Survey of Instructional Development Models* (4th ed.). Syracuse, NY: ERIC Clearinghouse on Information and Technology.

Haslam, B. (2010). *Teacher Professional Development Guide*. Washington, D.C.: The National Staff Development Council.

Haslam, B. (2008). *Maryland Teacher Professional Development Guide*. Hartford County (Maryland) Public Schools.

Hatch, M.J., & Cunliffe, A. L. (2013). *Organization Theory: Modern, Symbolic, and Postmodern Perspectives* (3rd ed*).* Oxford, UK: Oxford University Press.

Harding, J., Fox, C., & Mehta, J. (2002). Studying rare events through qualitative case studies: Lessons from a study of rampage school shootings. *Sociological Methods & Research*, *31*(2), 174 - 217.

Hawley, W.D., & Valli, L. (1999). The essentials of effective professional development: A new consensus. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the Learning Profession: Handbook of Policy and Practice.* San Francisco, CA: Jossey-Bass.

Hendricks, C., (2013). *Improving Schools Through Action Research: A Reflective Practice Approach* (3rd ed.). Boston, MA: Pearson Publishing.

Henry, G. T., Julnes, G., & Mark, M. M. (1998). Realist evaluation: An emerging theory in support of practice. *New Directions for Evaluation, no 78.*

Hogan, R.L. (2007). The historical development of program evaluation: Exploring the past and present. *Online Journal of Workforce Education and Development*, *2*(4).

Howe, K. (2004). A critique of experimentalism. *Qualitative Inquiry, 10*(1), 42-61.

Ithaca City School District Instructional Coaching Program (2013). *Assessment of the Budget Funding Database with a Communication, Implementation, and Evaluation Plan.* Draft, Cornell Institute for Public Affairs CIPA Domestic Capstone.

Joint Committee on Standards for Educational Evaluation. (2011). *The Program Evaluation Standards* (3rd ed.). Thousand Oaks, CA: Sage.

Joyce, B., & Showers, B. (2002). *Student Achievement Through Staff Development* (3rd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.

Kaufmann, P. & Keller, J. M. (1994) Levels of evaluation: Beyond Kirkpatrick. *HRD Quarterly, 5*(4), 371-380.

Kieser, A., & Ebers, M. (2014). *Organisationstheorien (*7th ed.). Stuttgart, Deutschland: Kohlhammer.

Killion, J., Harrison, C., Bryan, C., & Clifton, H. (2012). *Coaching Matters.* Oxford, OH: Learning Forward.

Killion, J. (2002). *Assessing Impact: Evaluating Staff Development*. Oxford, OH: National Staff Development.

Kirkpatrick, D. L. (1959). Techniques for evaluating training programs. *Journal of American Society of Training Directors, 13*(3), 21-26.

Knight, J. (2007*). Instructional Coaching: A Partnership Approach to Improving Instruction*. Thousand Oaks, CA: Corwin Press.

Knowles, M.S., Holton, E.F., & Swanson, R.A. (2005). *The Adult Learner: The Definite Classic in Adult Education and Human Resource Development* (6th ed.). Woburn, MA: Butterworth-Heinemann.

Laszlo, A., & Krippner, S. (1998). Systems theories: Their origins, foundations, and development. In J. S. Jordan (Ed). *Systems Theories and A Priori Aspects of Perception.* Amsterdam: Elsevier Science.

LCM mathleadership.com (2011). Instructional Implementation Sequence: Attaining the CCSS Mathematical Practices Engagement Strategies.

Loucks-Horsley, S., Hewson, P.W., Love, N. & Stiles, K. E. (1998). *Designing Professional Development for Teachers of Science and Mathematics.* Thousand Oaks, CA: Corwin Press.

Mark, M. M., Julnes, G., & Henry, G. T. (2000). *Evaluation: An Integrated Framework for Understanding, Guiding, and Improving Policies and Program.* San Francisco, CA: Jossey-Bass.

Marsh, J. A., Sloan McCombs, J., Lockwood, J. R., Martorell, F., Gershwin, D., Naftel, S., et al. (2008). *Supporting Literacy Across the Sunshine State: A Study of Florida Middle School Reading Coaches.* Santa Monica, CA: RAND.

Maxwell, J. A. (2004). Causal explanation, qualitative research, and scientific inquiry in education. *Educational Researcher, 33*(2), 3-11.

Maxwell, J. A. (2012). *A Realist Approach for Qualitative Research*. Thousand Oaks, CA: Sage.

Maxwell, J. A. (2013): *Qualitative Research Design: An Interactive Approach* (3rd ed.). Thousand Oaks, CA: Sage.

Maxwell, J. A., & Mittapalli, K. (2010). Realism as a stance for mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *SAGE Handbook of Mixed Methods in Social & Behavioral Research* (2nd ed.). Thousand Oaks, CA: Sage.

McLaughlin, J. A., & Jordan, G. B. (2010). Using logic models. In J. S. Wholey, H. P. Hatry, & K. E. Newcomer (Eds). *Handbook of Practical Performance Evaluation* (3rd ed.). San Francisco, CA: Jossey-Bass.

Merola, S., Horwood, T., Nebbergall, A., Uekawa, K., Passa, K., Mack, A., & Sun, J. (2011). *Evaluation of the Mathematics Instructional Coaches Pilot Program: A High School Success Pilot Program*, February 2011 Report. Fairfax, VA: ICF International.

Molenda, M., & Russell, J.D. (2006). Instruction as an intervention. In J. A. Pershing (Ed). *Handbook of Human Performance Technology: Principles, Practices, and Potential* (3rd ed.). San Francisco, CA: Pfeiffer.

Neuman, W. (2006): *Social Research Methods: Qualitative and Quantitative Approaches.* Boston, MA: Allyn & Bacon.

Newcomer, K. E., Hatry, H. P., & Wholey, J. S. (2010). Planning and designing useful evaluations. In J. S. Wholey, H. P. Hatry, & K. E. Newcomer (Eds). *Handbook of Practical Performance Evaluation* (3rd ed.). San Francisco, CA: Jossey-Bass.

Norton, L. S. (2009). *Action Research in Teaching and Learning*. New York, NY: Routledge.

O'Connell, J. F. (2009). *A Model for Assessing the Effectiveness of Professional Development for Improving student learning.* Dissertation Iowa State University, Ames, IW.

Patton, M. Q. (2001) *Qualitative Research and Evaluation Methods* (3rd ed.) Thousand Oaks: Sage.

Patton, M. Q. (2008). *Utilization-Focused Evaluation (*4th ed*).* Thousand Oaks: Sage.

Pawson R., Tilley, N. (1997). *Realistic Evaluation.* London: Sage.

Pawson, R., & Tilley, N. (2004). *Realist Evaluation*. White Paper. British Cabinet Office, London.

Pawson R. (2006). *Evidence-Based Policy: A Realist Perspective.* London: Sage.

Pawson R. (2013). *The Science of Evaluation: A Realist Manifesto*. London: Sage.

Penuel, W. R., Fishman, B., Yamaguchi, R., & Gallagher, L.P. (2007). What makes

professional development effective? Strategies that foster curriculum implementation. *American Educational Research Journal*, *44*(4), 921-958.

Perkins, D. N., & Salomon, G. (1987). Transfer and teaching thinking. In Perkins D. N., Lochhead, J., Bishop, J. (eds) *Thinking: The Second International Conference.* Hillsdale, NJ: Earlbaum.

Pershing, J.A. (2006). Human Performance Technology fundamentals. In J. A. Pershing (Ed). *Handbook of Human Performance Technology: Principles, Practices, and Potential* (3rd ed.). San Francisco, CA: Pfeiffer.

Pershing J. A. (Ed) (2006). *Handbook of Human Performance Technology: Principles, Practices, and Potential* (3rd ed.). San Francisco, CA: Pfeiffer.

Pfeffer, J., & Sutton, R.L. (2000). *The Knowing-Doing Gap: How Smart Companies Turn Knowledge into Action*. Boston, MA: Harvard Business School Press.

Phillips, J. J. (1997). *Training Evaluation and Measurement Methods: Proven Models and Methods for Evaluating HRD Programs* (3rd ed.). Woburn, MA: Butterworth-Heinemann

Posavac, E. (2011). *Program Evaluation.* Upper Saddle River, NJ: Pearson.

Reeves, D.B. (2010). *Transforming Professional development into Student Results.* Alexandria, VA: Association for Supervision and Curriculum Development.

Reiser, R.A. (2001). A history of instructional design and technology: Part II: A history of instructional design. *Educational Technology Research and Development*, *49*(2), 57-67.

Rogers, A., Petrosino, T., Huebner, T. A., & Hasci, T. A. (2000). Program theory evaluation: Practice, promise, and problems. *New Directions for Evaluation*, *87*(4), 5-12.

Ross, A. H. (2010). *Application and Utility of the Guskey Professional Development Evaluation Model in a Community College Setting.* Dissertation University of Tennessee, Knoxville, TE.

Scriven, M. (1998). Minimalist theory: The least theory that practice requires. *American Journal of Evaluation*, 19(1), 57-70.

Schein, E. H. (2010). *Organizational Culture and Leadership* (4th ed.). San Francisco, CA: Jossey-Bass.

Senge, P. M. (2006). *The Fifth Discipline: The Art and Practice of the Learning Organization* (rev. ed.). New York, NY: Doubleday Currency.

Showers, B., Joyce, B., Bennett, B. (1987). Synthesis of research on staff development: A framework for future study and state-of-the art analysis. *Educational Leadership, 45*(3), 77-87.

Smith, A. (1776/1993). *An Inquiry into the Nature and Causes of the Wealth of Nations*.

Spaulding D.T., & Smith, G. (2012). *Instructional Coaches and the Instructional Leadership Team.* Thousand Oaks, CA: Corwin Press.

Stamford Public Schools (2014a). Classroom Embedded Professional development Protocol.

Stamford Public Schools (2014b). Alliance District Improvement Plan (ADIP).

Stame, N. (2004). Theory-based evaluation and types of complexity. *Evaluation, 10*(1), 58-76.

State Education Resource Center (2007). The Connecticut Plan - School Reform. Draft. http://www.sde.ct.gov/sde/lib/sde/pdf/pressroom/TheConnecticutPlan.pdf

Stolovitch, H.D. (2000). Human Performance Technology: Research and theory to practice. *Performance Improvement*, 39(4), 7-16

Tallerico, M. (2005). *Supporting and Sustaining Teachers' Professional Development: A Principal's Guide.* Thousand Oaks, CA: Corwin Press.

Taylor-Powell, E., Steele, S., & Douglah, M. (1996). *Planning a Program Evaluation*. Madison, WI: University of Wisconsin - Cooperative Extension.

Wei, R. C., Darling-Hammond, L., Andree, A., Richardson, N., Orphanos, S. (2009). *Professional learning in the learning profession: A status report on teacher development in the United States and abroad.* Dallas, TX. National Staff Development Council.

Williams, B., Brown, T., & Onsman, A. (2010). Exploratory factor analysis: A five-step guide for novices. *Australasian Journal of Paramedicine*, 8(3). Retrieved from http://ro.ecu.edu.au/jephc/vol8/iss3/1

Wholey, J. S., Hatry, H. P., & Newcomer, K. E. (Eds.) (2010). *Handbook of Practical Program Evaluation* (3rd ed.). San Francisco, CA: Jossey-Bass.

Wilson, S.M., & Berne, J. (1999). Teacher learning and the acquisition of professional knowledge: An examination of research on contemporary professional development. *Review of Research in Education*, 24, 173-209.

W.K. Kellogg Foundation (2004a). *Evaluation Handbook*. Battle Creek, MI.

W.K. Kellogg Foundation (2004b). *Logic Model Development Guide.* Battle Creek, MI.

Vescio, V., Ross, D.,& Adams, A. (2008). A review of research on the impact of professional learning communities on teaching practice and student learning. *Teaching and Teacher Education*, 24, 80-91.

Yin, R. K. (2009). *Case study research: Design and methods* (4th ed.). Thousand Oaks, CA: Sage.

Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., & Shapley, K. (2007). Reviewing the evidence on how teacher professional development affects student achievement. *Issues & Answers Report*, REL 2007–No. 033.U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest.